Asymptotic moments of spatial branching processes

Isaac Gonzalez[1] · Emma Horton[2] · Andreas E. Kyprianou[1]

Abstract

Suppose that $X = (X_t, t \geq 0)$ is either a superprocess or a branching Markov process on a general space E, with non-local branching mechanism and probabilities $\mathbb{P}_{\delta_x}$, when issued from a unit mass at $x \in E$. For a general setting in which the first moment semigroup of X displays a Perron–Frobenius type behaviour, we show that, for $k \geq 2$ and any positive bounded measurable function f on E,

$$\lim_{t \to \infty} g_k(t) \mathbb{E}_{\delta_x}[\langle f, X_t \rangle^k] = C_k(x, f),$$

where the constant $C_k(x, f)$ can be identified in terms of the principal right eigenfunction and left eigenmeasure and $g_k(t)$ is an appropriate deterministic normalisation, which can be identified explicitly as either polynomial in t or exponential in t, depending on whether X is a critical, supercritical or subcritical process. The method we employ is extremely robust and we are able to extract similarly precise results that additionally give us the moment growth with time of $\int_0^t \langle f, X_s \rangle \mathrm{d}s$, for bounded measurable f on E.

Keywords Moments · Branching processes · Superprocesses · Non-local branching · Asymptotic behaviour

Mathematics Subject Classification 60J68 · 60J80

✉ Andreas E. Kyprianou
a.kyprianou@bath.ac.uk

Isaac Gonzalez
igg22@bath.ac.uk

Emma Horton
emma.horton@inria.fr

1 Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK

2 INRIA, Bordeaux Research Centre, 33405 Talence, France

Springer

Contents

1 Introduction and main results

A fundamental question concerning general spatial branching processes, both super-processes and branching Markov processes, pertains to their moments. Whilst the setting of first and second moments has received quite some attention, limited information seems to be known about higher moments, in particular, their asymptotic behaviour with time. Relevant references that touch upon this topic include [12, 17, 18, 21, 24]. In this paper, we provide a single general result that pertains to both super-processes and spatial branching Markov processes and which provides a very precise and somewhat remarkable result for moment growth.

We show that, under the assumption that the first moment semigroup of the process exhibits a natural Perron Frobenious type behaviour, the k-th moment functional of either a superprocess or branching Markov process, when appropriately normalised, limits to a precise constant. The setting in which we work is remarkably general, even allowing for the setting of non-local branching; that is, where mass is created at a different point in space to the position of the parent. Moreover, the methodology we use appears to be extremely robust and we show that the asymptotic k-th moments of the running occupation measure are equally accessible using essentially the same approach. Our results will thus expand on what is known for branching diffusions and superdiffusions e.g. in [10], [23], as well as giving precise growth rates for the moments of occupations. In future work we hope to use the ideas in this paper to develop general central limit theorems for the aforesaid class of processes.

To this end, let us spend some time providing the general setting in which we wish to work. Let E be a Lusin space. Throughout, will write $B(E)$ for the Banach space of bounded measurable functions on E with norm $\|\cdot\|$, $B^+(E)$ for non-negative bounded

measurable functions on E and $B_1^+(E)$ for the subset of functions in $B^+(E)$ which are uniformly bounded by unity. We are interested in spatial branching processes that are defined in terms of a Markov process and a branching operator. The former can be characterised by a semigroup on E, denoted by $\mathtt{P} = (\mathtt{P}_t, t \geq 0)$. We do not need $\mathtt{P}$ to have the Feller property, and it is not necessary that $\mathtt{P}$ is conservative. That said, if so desired, we can append a cemetery state $\{\dagger\}$ to E, which is to be treated as an absorbing state, and regard $\mathtt{P}$ as conservative on the extended space $E \cup \{\dagger\}$, which can also be treated as a Lusin space. Equally, we can extend the branching operator to $E \cup \{\dagger\}$ by defining it to be zero on $\{\dagger\}$, i.e. no branching activity on the cemetery state.

1.1 Branching Markov processes

Consider now a spatial branching process in which, given their point of creation, particles evolve independently according to a $\mathtt{P}$-Markov process. In an event which we refer to as 'branching', particles positioned at x die at rate $\beta(x)$, where $\beta \in B^+(E)$, and instantaneously, new particles are created in E according to a point process. The configurations of these offspring are described by the random counting measure

$$\mathcal{Z}(A) = \sum_{i=1}^{N} \delta_{x_i}(A),$$

for Borel A in E. The law of the aforementioned point process depends on x, the point of death of the parent, and we denote it by $\mathcal{P}_x$, $x \in E$, with associated expectation operator given by $\mathcal{E}_x$, $x \in E$. This information is captured in the so-called branching mechanism

$$\mathtt{G}[f](x) := \beta(x)\mathcal{E}_x\left[\prod_{i=1}^{N} f(x_i) - f(x)\right], \qquad x \in E, \tag{1}$$

where we recall $f \in B_1^+(E) := \{f \in B^+(E) : \sup_{x \in E} f(x) \leq 1\}$. Without loss of generality we can assume that $\mathcal{P}_x(N = 1) = 0$ for all $x \in E$ by viewing a branching event with one offspring as an extra jump in the motion. On the other hand, we do allow for the possibility that $\mathcal{P}_x(N = 0) > 0$ for some or all $x \in E$.

Henceforth we refer to this spatial branching process as a $(\mathtt{P}, \mathtt{G})$-branching Markov process. It is well known that if the configuration of particles at time t is denoted by $\{x_1(t), \ldots, x_{N_t}(t)\}$, then, on the event that the process has not become extinct or exploded, the branching Markov process can be described as the co-ordinate process $X = (X_t, t \geq 0)$ in the space of atomic measures on E with non-negative integer total mass, denoted by $N(E)$, where

$$X_t(\cdot) = \sum_{i=1}^{N_t} \delta_{x_i(t)}(\cdot), \qquad t \geq 0.$$

In particular, X is Markovian in $N(E)$. Its probabilities will be denoted $\mathbb{P} := (\mathbb{P}_\mu, \mu \in N(E))$. With this notation in hand, it is worth noting that the independence that is manifest in the definition of branching events and movement implies that if we define,

$$v_t[f](x) = \mathbb{E}_{\delta_x}\left[\prod_{i=1}^{N_t} f(x_i(t))\right], \qquad f \in B_1^+(E), t \geq 0, \tag{2}$$

then for $\mu \in N(E)$ given by $\mu = \sum_{i=1}^n \delta_{y_i}$, we have

$$\mathbb{E}_\mu\left[\prod_{i=1}^{N_t} f(x_i(t))\right] = \prod_{i=1}^n v_t[f](y_i), \qquad t \geq 0. \tag{3}$$

Moreover, for $f \in B^+(E)$ and $x \in E$,

$$v_t[f](x) = \hat{\mathsf{P}}_t[f](x) + \int_0^t \mathsf{P}_s\left[\mathsf{G}[v_{t-s}[f]]\right](x)\mathrm{d}s, \qquad t \geq 0, \tag{4}$$

where $\hat{\mathsf{P}}_t$ is defined similarly to P_t but returns a value of 1 on the event of killing. The above equation describes the evolution of the semigroup v_t: either the initial particle has not branched (and has possibly been absorbed) by time t or at some time $s \leq t$, the initial particle has branched, producing offspring according to G. We refer the reader to [19, 22] for a proof.

Branching Markov processes enjoy a very long history in the literature, dating back as far as [30–32], with a broad base of literature that is arguably too voluminous to give a fair summary of here. Most literature focuses on the setting of local branching. This corresponds to the setting that all offspring are positioned at their parent's point of death (i.e. $x_i = x$ in the definition of G). In that case, the branching mechanism reduces to

$$\mathsf{G}[s](x) = \beta(x)\left[\sum_{k=0}^\infty p_k(x)s^k - s\right], \qquad x \in E,$$

where $s \in [0, 1]$ and $(p_k(x), k \geq 0)$ is the offspring distribution when a parent branches at site $x \in E$. The branching mechanism G may otherwise be seen in general as a mixture of local and non-local branching.

1.2 Superprocesses

Superprocesses can be thought of as the high-density limit of a sequence of branching Markov processes, resulting in a new family of measure-valued Markov processes; see e.g. [6, 7, 13, 26, 34]. Just as branching Markov processes are Markovian in $N(E)$, the former are Markovian in the space of finite Borel measures on E topologised by the weak convergence topology, denoted by $M(E)$. There is a broad literature base for

superprocesses, e.g. [6, 15, 17, 26, 34], with so-called local branching mechanisms, and later broadened to the more general setting of non-local branching mechanisms in [7, 26]. Let us now introduce these concepts with an autonomous definition of what we mean by a superprocess.

A Markov process $X := (X_t : t \geq 0)$ with state space $M(E)$ and probabilities $\mathbb{P} := (\mathbb{P}_\mu, \mu \in M(E))$ is called a $(\mathbb{P}, \psi, \phi)$-superprocess if it has transition semigroup $(\hat{\mathbf{E}}_t, t \geq 0)$ on $M(E)$ satisfying

$$\mathbb{E}_\mu \left[e^{-\langle f, X_t \rangle} \right] = \int_{M(E)} e^{-\langle f, v \rangle} \hat{\mathbf{E}}_t(\mu, dv) = e^{-\langle V_t[f], \mu \rangle}, \qquad \mu \in M(E), f \in B^+(E). \tag{5}$$

Here, we work with the inner product on $B^+(E) \times M(E)$ defined by $\langle f, \mu \rangle = \int_E f(x) \mu(dx)$ and $(V_t, t \geq 0)$ is a semigroup evolution that is characterised via the unique bounded positive solution to the evolution equation

$$V_t[f](x) = P_t[f](x) - \int_0^t P_s \left[\psi(\cdot, V_{t-s}[f](\cdot)) + \phi(\cdot, V_{t-s}[f]) \right](x) ds. \tag{6}$$

Here ψ denotes the local branching mechanism

$$\psi(x, \lambda) = -b(x)\lambda + c(x)\lambda^2 + \int_{(0,\infty)} \left(e^{-\lambda y} - 1 + \lambda y \right) v(x, dy), \quad \lambda \geq 0, \tag{7}$$

where $b \in B(E)$, $c \in B^+(E)$ and $(x \wedge x^2)v(x, dy)$ is a bounded kernel from E to $(0, \infty)$, and ϕ is the non-local branching mechanism

$$\phi(x, f) = \beta(x) \left(f(x) - \zeta(x, f) \right), \tag{8}$$

where $\beta \in B^+(E)$ and ζ has representation

$$\zeta(x, f) = \gamma(x, f) + \int_{M(E)^\circ} (1 - e^{-\langle f, v \rangle}) \Gamma(x, dv), \tag{9}$$

such that $\gamma(x, f)$ is a bounded function on $E \times B^+(E)$ and $v(1)\Gamma(x, dv)$ is a bounded kernel from E to $M(E)^\circ := M(E) \setminus \{0\}$ with

$$\gamma(x, f) + \int_{M(E)^\circ} \langle 1, v \rangle \, \Gamma(x, dv) \leq 1. \tag{10}$$

We refer the reader to [7, 27] for more details regarding the above formulae. Lemma 3.1 in [7] tells us that the functional $\zeta(x, f)$ has the following equivalent representation

$$\zeta(x, f) = \int_{M_0(E)} \left[\gamma(x, \pi) \langle f, \pi \rangle + \int_0^\infty \left(1 - e^{-u\langle f, \pi \rangle} \right) n(x, \pi, du) \right] G(x, d\pi), \tag{11}$$

where $M_0(E)$ denotes the set of probability measures on E, $\gamma \geq 0$ is a bounded function on $E \times M_0(E)$, $un(x, \pi, \mathrm{d}u)$ is a bounded kernel from $E \times M_0(E)$ to $(0, \infty)$ and $G(x, \mathrm{d}\pi)$ is a probability kernel from E to $M_0(E)$ with

$$\gamma(x, \pi) + \int_0^\infty un(x, \pi, \mathrm{d}u) \leq 1. \tag{12}$$

The reader will note that we have deliberately used some of the same notation for both branching Markov processes and superprocesses. In the sequel there should be no confusion and the motivation for this choice of repeated notation is that our main result is indifferent to which of the two processes we are talking about.

1.3 Main results: k-th moments

As alluded to above, in what follows, $(X, \mathbb{P})$ is taken as either a branching Markov process or a superprocess as defined in the previous section. Our main results concern understanding the growth of the k-th moment functional in time

$$\mathrm{T}_t^{(k)}[f](x) := \mathbb{E}_{\delta_x}\left[\langle f, X_t \rangle^k\right], \qquad x \in E, f \in B^+(E), k \geq 1, t \geq 0. \tag{13}$$

For convenience, we will write T in preference of $\mathrm{T}^{(1)}$ throughout.

Before stating our main theorem, we first introduce some assumptions that will be crucial in analysing the moments defined above. First, we have a Perron–Frobenius-type assumption.

(H1): There exists an eigenvalue $\lambda \in \mathbb{R}$ and a corresponding right eigenfunction $\varphi \in B^+(E)$ and finite left eigenmeasure $\tilde{\varphi}$ such that, for $f \in B^+(E)$,

$$\langle \mathrm{T}_t[\varphi], \mu \rangle = \mathrm{e}^{\lambda t} \langle \varphi, \mu \rangle \text{ and } \langle \mathrm{T}_t[f], \tilde{\varphi} \rangle = \mathrm{e}^{\lambda t} \langle f, \tilde{\varphi} \rangle,$$

for all $\mu \in N(E)$ (resp. $M(E)$) if $(X, \mathbb{P})$ is a branching Markov process (resp. a superprocess). Further let us define

$$\Delta_t = \sup_{x \in E, f \in B_1^+(E)} |\varphi(x)^{-1} \mathrm{e}^{-\lambda t} \mathrm{T}_t[f](x) - \langle \tilde{\varphi}, f \rangle|, \qquad t \geq 0.$$

We suppose that

$$\sup_{t \geq 0} \Delta_t < \infty \text{ and } \lim_{t \to \infty} \Delta_t = 0. \tag{14}$$

Before explaining the heuristics behind (H1), we state are second assumption, which is a moment condition on the offspring distribution.

(H2): Suppose $k \geq 1$. If $(X, \mathbb{P})$ is a branching Markov process,

$$\sup_{x \in E} \mathcal{E}_x(\langle 1, \mathcal{Z} \rangle^k) < \infty \tag{15}$$

and if $(X, \mathbb{P})$ is a superprocess,

$$\sup_{x \in E} \left(\int_0^\infty |y|^k \nu(x, \mathrm{d}y) + \int_{M(E)^\circ} \langle 1, \nu \rangle^k \Gamma(x, \mathrm{d}\nu) \right) < \infty. \tag{16}$$

Let us spend a little time considering these two assumptions in more detail. For a lot of literature surrounding spatial branching processes, there has been emphasis on results for which an underlying assumption of exponential ergodic growth in the first moment is present as in (H1); see e.g. [1, 16, 19, 22, 27, 29]. Due to this, we may characterise the process as supercritical if $\lambda > 0$, critical if $\lambda = 0$ and subcritical if $\lambda < 0$.

One way to understand (14), is through the martingale that comes hand-in-hand with the eigenpair (λ, φ), i.e.

$$M_t^\varphi := \mathrm{e}^{-\lambda t} \langle \varphi, X_t \rangle, \qquad t \geq 0. \tag{17}$$

Normalising this martingale and using it as a change of measure results in the ubiquitous spine decomposition; cf. [21, 22, 29]. Roughly speaking, under the change of measure, the process is equal in law to a copy of the original process with a superimposed process of immigration, which occurs both in space and time along the path of a single particle trajectory in E, the spine. Moreover, the assumption (14) implies that the spine has a stationary limit with stationary measure $\varphi\tilde{\varphi}$.

The assumptions (15) and (16) of (H2) are natural to ensure that k-moments are well defined for all $t \geq 0$. If not explicitly stated in the literature, their need to ensure that the functional moments $\mathrm{T}_t^{(k)}[f](x)$ are finite for all $t \geq 0$, $f \in B^+(E)$ and $x \in E$ is certainly folklore. The two conditions (15) and (16) are clearly natural analogues of one another. Indeed, whereas for superprocesses, it is usual to separate out the non-diffusive local branching behaviour from non-local behaviour, i.e via the measures $\nu(x, \mathrm{d}y)$ and $\Gamma(x, \mathrm{d}\nu)$, the analogous behaviour is captured in the single point process $\mathcal{Z}$ for branching Markov processes.

In terms of the eigenvalue in (H1), the following suite of results give us the precise growth rates for k-th moments in each of the critical ($\lambda = 0$), supercritical ($\lambda > 0$) and subcritical ($\lambda < 0$) settings. In all three results, $(X, \mathbb{P})$ is either a (P, G)-branching particle system or a (P, ψ, ϕ)-superprocess on E.

Theorem 1 (Critical, $\lambda = 0$) *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda = 0$. Define*

$$\Delta_t^{(\ell)} = \sup_{x \in E, f \in B_1^+(E)} \left| t^{-(\ell-1)} \varphi(x)^{-1} T_t^{(\ell)}[f](x) - 2^{-(\ell-1)} \ell! \, \langle f, \tilde{\varphi} \rangle^\ell \, \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{\ell-1} \right|,$$

where

$$\mathbb{V}[\varphi](x) = \beta(x)\mathcal{E}_x \left(\langle \varphi, \mathcal{Z} \rangle^2 - \langle \varphi^2, \mathcal{Z} \rangle \right),$$

if $(X, \mathbb{P})$ *is a branching Markov process or*

$$\mathbb{V}[\varphi](x) = \psi''(x, 0+)\varphi(x)^2 + \beta(x) \int_{M(E)^\circ} \langle \varphi, \nu \rangle^2 \Gamma(x, \mathrm{d}\nu)$$

if $(X, \mathbb{P})$ *is a superprocess. Then, for all* $\ell \leq k$

$$\sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \to \infty} \Delta_t^{(\ell)} = 0. \tag{18}$$

The novel contribution of Theorem 1 is the fact that no such result currently exists in the literature at this level of generality. The only comparable results appear in [11], which uses similar methods to derive the convergence for the critical Crump-Mode-Jagers process[1] and in [20], which inspired this paper but only deals with the special case of a general *critical* branching particle processes where the test function f is specifically taken to be the eigenfunction φ. We will discuss these examples in more detail after stating our main results. We have been unable to find comparable results for superprocesses.

There are two facts that stand out in Theorem 1. The first is the polynomial scaling, which is quite a delicate conclusion given that there is no exponential growth to rely on. The second is that, for $k \geq 3$, the scaled moment limit is expressed not in terms of the k-th moments in (15) and (16), but rather the second order moments.

In some sense, however, both the polynomial growth and the nature of the limiting constant are not entirely surprising given the folklore for the critical setting. More precisely, in at least some settings (see e.g. [20]), one would expect to see a Yaglom-type result at criticality. The latter would classically see, conditional on survival, convergence in law of $t^{-1}\langle f, X_t \rangle$ to an exponentially distributed random variable as $t \to \infty$, whose parameter is entirely determined by the second moment of X. This implies that, conditional on survival, the limit of the k-th moment of $\langle f, X_t \rangle$ behaves like $k! t^k c^k$ (i.e. the k-th moment of the aforesaid exponential distribution) as $t \to \infty$, where $c > 0$ is written in terms of a second moment functional of the spatial offspring distribution. One can also expect a Kolmogorov-type result for the survival probability, which says that

$$\mathbb{P}_{\delta_x}(N_t > 0) \sim c\varphi(x)t^{-1}, \qquad \text{as } t \to \infty, \tag{19}$$

where we recall that $N_t = \langle 1, X_t \rangle$ is the total mass of the process at time t; see for example [20] for the particle system setting and [29] in the superprocess setting. Combining this with the moment asymptotic of

$$\mathrm{T}_t^{(k)}[f](x) = \mathbb{E}_{\delta_x}[\langle f, X_t \rangle^k | N_t > 0]\mathbb{P}_{\delta_x}(N_t > 0) \tag{20}$$

implies that the k-th moment behaves like $\varphi(x)t^{k-1}k!c^k$, which is precisely the result given in Theorem 1.

---

[1] The article [11] only came to light after submitting this article.

It is also curious why, at criticality, all higher moments can be expressed in terms of the constant c, which is written in terms of a second moment functional of the spatial offspring distribution. We can give some intuition here, at least in the branching Markov setting, as to why this is the case. Let us recall the classical folklore that conditioning on survival is equivalent to studying the process under the change of measure

$$\left.\frac{\mathrm{d}\mathbb{P}^{\varphi}_{\delta_x}}{\mathrm{d}\mathbb{P}_{\delta_x}}\right|_{\sigma(X_s, s \leq t)} = \frac{\langle \varphi, X_t \rangle}{\varphi(x)}, \qquad x \in E, t \geq 0, \tag{21}$$

The above change of measure (21) yields the classical *spine decomposition*, cf. [20]. In particular, this means that we can write, for $f \in B^+(E)$,

$$\langle f, X_t \rangle = f(\xi_t) + \sum_{i=1}^{n_t} \sum_{j=1}^{N_i} \langle f, X^{i,j}_{t-T_i} \rangle, \tag{22}$$

where $\xi = (\xi_t, t \geq 0)$ is the motion of an immortal particle, called the *spine*, whose semigroup is conservative, n_t is the number of branching events along the spine, which arrive at a rate which depends on the motion of ξ, N_i is the number of offspring produced such that, at the i-th such branching event, which occurs at time $T_i \leq t$, $X^{i,j}_{t-T_i}, j = 1, \ldots, N_i$ are i.i.d copies of the original branching Markov process under $\mathbb{P}_{\xi_{T_i}}$, which provide mass at time t. In other words, this means that under $\mathbb{P}^{\varphi}$, the process X can be decomposed into a single immortal trajectory, off which, copies of the original process $(X, \mathbb{P})$ immigrate simultaneously to groups of siblings.

With this in mind, let us consider genealogical lines of descent that contribute to the bulk of the mass of the k-th moment at large times t. For each copy of $(X, \mathbb{P})$ that immigrates onto the spine at time $s > 0$, the probability that the process survives to time $t \geq s$, thus contributing to the bulk of the k-th moment at time t, is $O(1/(t-s)) \approx O(1/t)$; cf. (19). If there are multiple offspring at an immigration event at time s, then the chance that at least two of these offspring contribute to the bulk of the k-th moment at time t is $O(1/t^2)$. Moreover, the semigroup of the spine ξ is given by $\mathrm{P}^{\varphi}_t[f](x) = \mathrm{P}_t[\varphi f](x)/\varphi(x)$, which, under (H1), limits to a stationary distribution $\varphi(x)\tilde{\varphi}(\mathrm{d}x), x \in E$. This has the effect that the arrival of branching events along the spine begin to look increasingly like a Poisson process as $t \to \infty$. Hence, for large t, $n_t \approx O(t)$.

Putting these pieces together, as $t \to \infty$, there are approximately $O(t)$ branch points along the spine, each of which has the greatest likelihood of a single offspring among immigrating siblings contributing to the bulk of the k-th moment at time t, with probability of order $O(1/t)$. Thus, it is clear that we only expect to see one of each sibling group of immigrants along the spine contributing to the mass of the k-th moment at time t. Now let β^{φ} denote the spatial rate at which offspring immigrates onto the spine and let $\{x_1, \ldots, x_N\}$ denote their positions at the point of branching *including* the position of the spine at this instance. Let $\mathcal{P}^{\varphi}$ denote the law of this offspring distribution, and suppose that i^* is the (random) index of the offspring that continues the evolution of the spine. We have that the rate at which a 'uniform selection' of a

single offspring which is not the spine at a branching event (seen through the function $f \in B^+(E)$) is given by

$$\beta^\varphi(x)\mathcal{E}_x^\varphi\left[\sum_{i=1}^N f(x_i)\mathbf{1}_{(i \neq i^*)}\right]$$

$$= \beta(x)\frac{\mathcal{E}_x\left[\langle\varphi, \mathcal{Z}\rangle\right]}{\varphi(x)}\mathcal{E}_x\left[\frac{\langle\varphi, \mathcal{Z}\rangle}{\mathcal{E}_x\left[\langle\varphi, \mathcal{Z}\rangle\right]}\sum_{i=1}^N\frac{\varphi(x_i)}{\langle\varphi, \mathcal{Z}\rangle}\sum_{\substack{i=1 \\ j \neq i}}^N f(x_j)\right]$$

$$= \frac{\beta(x)}{\varphi(x)}\mathcal{E}_x\left[\sum_{i=1}^N\varphi(x_i)\sum_{\substack{i=1 \\ j \neq i}}^N f(x_j)\right]$$

$$= \frac{\beta(x)}{\varphi(x)}\mathcal{E}_x\left[\langle f, \mathcal{Z}\rangle\langle\varphi, \mathcal{Z}\rangle - \langle\varphi f, \mathcal{Z}\rangle\right], \tag{23}$$

where we have used, from [20], that $\beta^\varphi(x) = \beta(x)\mathcal{E}_x\left[\langle\varphi, \mathcal{Z}\rangle\right]/\varphi(x)$, $\mathcal{P}_x^\varphi$ is absolutely continuous with respect to $\mathcal{P}_x$ with density $\langle\varphi, \mathcal{Z}\rangle/\mathcal{E}_x\left[\langle\varphi, \mathcal{Z}\rangle\right]$ and, given $\{x_1, \ldots, x_N\}$, $i^* = i$ is empirically selected with probability proportional to $\varphi(x_i)$.

We know from the the setting of first moments that it is the projection of $\langle f, X_t\rangle$ on to $\langle\varphi, X_t\rangle$, with coefficient $\langle f, \tilde\varphi\rangle$, which dominates the mean growth; indeed a strong law large numbers exists to this effect as well, see e.g. [19]. In this spirit, let us take $f = \varphi$ for simplicity and we see that in (23) we get precisely $\mathbb{V}[\varphi](x)/\varphi(x)$ on the right-hand side. Hence, finally, we conclude our heuristic by observing that the rate at which immigration off the spine contributes to the bulk of the k-th moment limit of $\langle\varphi, X_t\rangle$ is determined by the second moment functional $\mathbb{V}[\varphi]$; together with (20) and the associated remarks above, this goes some way to explaining the appearance of the limit in Theorem 1.

The next results present a significantly different picture for the supercritical and subcritical cases. For those settings, the exponential behaviour of the first moment semigroup becomes a dominant feature of the higher moments.

Theorem 2 (Supercritical, $\lambda > 0$) *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda > 0$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x \in E, f \in B_1^+(E)}\left|\varphi(x)^{-1}\mathrm{e}^{-\ell\lambda t}T_t^{(\ell)}[f](x) - \ell!\langle f, \tilde\varphi\rangle^\ell L_\ell\right|,$$

where $L_1 = 1$ and we define iteratively for $k \geq 2$

$$L_k = \frac{1}{\lambda(k-1)}\left\langle\beta\mathcal{E}.\left[\sum_{[k_1,\ldots,k_N]_k^2}\prod_{\substack{j=1 \\ j:k_j>0}}^N\varphi(x_j)L_{k_j}\right], \tilde\varphi\right\rangle,$$

where $[k_1, \ldots, k_N]_k^2$ is the set of all non-negative N-tuples $(k_1, \ldots, k_N)$ such that $\sum_{i=1}^N k_i = k$ and at least two of the k_i are strictly positive[2] if $(X, \mathbb{P})$ is a branching Markov process, or

$$
L_k(x) = \sum_{\{m_1, \ldots, m_{k-1}\}_k} \frac{1}{m_1! \ldots m_{k-1}!} (m_1 + \cdots + m_{k-1} - 1)! \varphi(x)^{m_1 + \cdots + m_{k-1} - 1}
$$

$$
\prod_{j=1}^{k-1} \left(-L_j(x) \right)^{m_j} + \frac{1}{\lambda(k-1)} \langle \mathbb{V}_k [\varphi], \tilde{\varphi} \rangle
$$

and iteratively with $\mathbb{V}_2 [\varphi] (x) = \mathbb{V} [\varphi] (x)$ (defined in the previous theorem) and for $k \geq 3$

$$
\mathbb{V}_k[\varphi](x) = \sum_{\{m_1, \ldots, m_{k-1}\}_k} \frac{1}{m_1! \ldots m_{k-1}!} \prod_{j=2}^{k-1} \left(\frac{\langle \mathbb{V}_j [\varphi], \tilde{\varphi} \rangle}{\lambda(j-1)} \right)^{m_j}
$$

$$
\times \left[\psi^{(m_1 + \cdots + m_{k-1})}(x, 0+)(-\varphi(x))^{m_1 + \cdots + m_{k-1}} + \beta(x) \right.
$$

$$
\left. \int_{M(E)^\circ} \langle \varphi, v \rangle^{m_1 + \cdots + m_{k-1}} \Gamma(x, dv) \right]
$$

if $(X, \mathbb{P})$ is a superprocess. Here the sums run over the set $\{m_1, \ldots, m_{k-1}\}_k$ of positive integers such that $m_1 + 2m_2 + \cdots + (k-1)m_{k-1} = k$. Then, for all $\ell \leq k$

$$
\sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \quad \text{and} \quad \lim_{t \to \infty} \Delta_t^{(\ell)} = 0. \tag{24}
$$

Although we have not been able to find existing results of this kind in such generality for supercritical processes, we note that the asymptotics of branching diffusions with either constant or compactly supported branching potentials were studied in [25]. Moreover, the asymptotics for the first and second moments of age-dependent branching processes were considered in [4] and [5], respectively. While Jensen's inequality easily shows that this is the minimal rate of growth, it turns out that it is the exact rate of growth, implying that the k-th moment grows as the k-th power of the first moment, i.e. using the terminology of [35], there is no intermittency. If we again appeal to folklore then this is again not necessarily surprising. In a number of settings, we would expect X to obey a strong law of large numbers (cf. [1, 16, 19, 27]) in the sense that

$$
\lim_{t \to \infty} e^{-\lambda t} \langle f, X_t \rangle = \langle \tilde{\varphi}, f \rangle M_\infty^\varphi,
$$

where $(M_t^\varphi, t \geq 0)$ was defined in (17) and the limit holds either almost surely or in the sense of L^p moments, for $p > 1$. Moreover, returning to the heuristic involving the spine decomposition discussed after Theorem 1, in this case, the copies of the

---

[2] We interpret $\sum_\emptyset = 0$ and $\prod_\emptyset = 1$.

original process that branch off the spine are all supercritical so we expect them all to survive to time t and hence why we see the sum over tuples with at least two positive entries in the recurrence L_k (the argument for not seeing precisely one still holds).

It is also worth commenting on the difference in the dependence on x in the limit for the two processes. In the case of the branching Markov process, the dependence on x appears through the normalisation by $\varphi(x)$, which appears due to the assumption (H1). For superprocesses, this dependence is much more complicated. Thinking in terms of what is known as the *skeletal decomposition* (see e.g. [2, 14]), the superprocess issued from a unit mass at x can be seen at time $t > 0$ as the aggregation of a Poisson point process of 'superprocess excursions' conditioned to survive beyond time t as well as a copy of the superprocess conditioned to become extinct by time t. In the supercritical setting, a finite Poisson number of these excursions will contribute to the overall growth of the process, as $t \to \infty$, with a rate proportional to $p(x)\delta_x$, where $p(x)$ is rate of survival of an excursion issued from $x \in E$. Moreover, in terms of mass, each of the excursions is of the same order of magnitude as an analogous branching particle system. Taking k-th moments of the x-dependent Poisson sum of such excursions introduces an additional layer of complexity to its asymptotic behaviour and, specifically, the dependency on x. This may go part way to explaining the dependency of L_k on x in that setting.

Finally we turn to the growth of moments in the subcritical setting, which offers the heuristically appealing result that the k-th moment decays slower than the k-th moment of the linear semigroup.

Theorem 3 (Subcritical, $\lambda < 0$) *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda < 0$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} \mathrm{e}^{-\lambda t} T_t^{(\ell)}[f](x) - \ell! \langle f, \tilde{\varphi} \rangle^\ell L_\ell \right|,$$

where we define iteratively $L_1 = 1$ and for $k \geq 2$,

$$L_k = \frac{\langle f^k, \tilde{\varphi} \rangle}{\langle f, \tilde{\varphi} \rangle^k k!} + \left\langle \beta \mathcal{E}. \left[\sum_{n=2}^{k} \frac{1}{|\lambda|(n-1)} \sum_{[k_1,\ldots,k_N]_k^n} \prod_{\substack{j=1 \\ j:k_j>0}}^{N} \varphi(x_j) L_{k_j} \right], \tilde{\varphi} \right\rangle,$$

where $[k_1, \ldots, k_N]_k^n$ is the set of all non-negative N-tuples $(k_1, \ldots, k_N)$ such that $\sum_{i=1}^{N} k_i = k$ and exactly $2 \leq n \leq k$ of the k_i are strictly positive if $(X, \mathbb{P})$ is a branching Markov process, or $L_k = \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle$, where $\mathbb{V}_1[\varphi](x) = \varphi(x)$ and for $k \geq 2$,

$$\mathbb{V}_k[\varphi](x) = \sum_{\{m_1,\ldots,m_{k-1}\}_k} \frac{1}{m_1! \ldots m_{k-1}!} \frac{1}{\lambda(1 - m_1 - \cdots - m_{k-1})} \prod_{j=1}^{k-1} \langle \mathbb{V}_j[\varphi], \tilde{\varphi} \rangle^{m_j}$$
$$\times \left[\psi^{(m_1+\cdots+m_{k-1})}(x, 0+)(-\varphi(x))^{m_1+\cdots+m_{k-1}} \right.$$

$$+\beta(x) \int_{M(E)^\circ} \langle \varphi, v \rangle^{m_1 + \cdots + m_{k-1}} \Gamma(x, dv) \Bigg]$$

if $(X, \mathbb{P})$ is a superprocess. Here the sums run over the set $\{m_1, \ldots, m_{k-1}\}_k$ of non-negative integers such that $m_1 + 2m_2 + \cdots + (k-1)m_{k-1} = k$. Then, for all $\ell \leq k$

$$\sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \to \infty} \Delta_t^{(\ell)} = 0. \tag{25}$$

As alluded to above, it is heuristically appealing that the the k-th moment does not grow at the rate $\exp(-k\lambda t)$. On the other hand the actual growth rate $\exp(-\lambda t)$ is slightly less obvious but nonetheless the obvious candidate. The decay in mass to zero in the branching system would suggest that the k-th moment similarly does so, but no slower that the first moment.

1.4 Main result: *k*-th occupation moments

It also transpires that our method is remarkably robust. Indeed, again taking an agnostic position on whether X is a branching Markov process or a superprocess, as we will show, careful consideration of the proofs of Theorems 1, 2 and 3 demonstrate that we can also conclude results for the quantities

$$\mathrm{M}_t^{(k)}[g](x) := \mathbb{E}_{\delta_x} \left[\left(\int_0^t \langle g, X_s \rangle \, ds \right)^k \right], \quad x \in E, g \in B^+(E), k \geq 1, t \geq 0.$$

We can think of $\int_0^t \langle g, X_s \rangle \, ds$ as characterising the running occupation measure $\int_0^t X_s(\cdot) \, ds$ of the process X and hence we refer to $\mathrm{M}_t^{(k)}[g](x)$ as the k-th moment of the running occupation. The following results also emerge from our calculations, mirroring Theorems 1, 2 and 3 respectively.

Theorem 4 (Critical, $\lambda = 0$) *Suppose that (H1) holds along with (H2) for $k \geq 2$ and $\lambda = 0$. Define*

$$\Delta_t^{(\ell)} = \sup_{x \in E, g \in B_1^+(E)} \left| t^{-(2\ell-1)} \varphi(x)^{-1} \mathrm{M}_t^{(\ell)}[g](x) - 2^{-(\ell-1)} \ell! \, \langle g, \tilde{\varphi} \rangle^\ell \, \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{\ell-1} L_\ell \right|,$$

where $L_1 = 1$ and L_k is defined through the recursion $L_k = (\sum_{i=1}^{k-1} L_i L_{k-i})/(2k-1)$ if $(X, \mathbb{P})$ is a branching Markov process or $L_k = (\sum_{\{k_1, k_2\}^+} L_{k_1} L_{k_2})/(2k-1)$ where $\{k_1, k_2\}^+$ is the set of non-negative integers k_1, k_2 such that $k_1 + k_2 = k$ if $(X, \mathbb{P})$ is a superprocess. Then, for all $\ell \leq k$

$$\sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \to \infty} \Delta_t^{(\ell)} = 0. \tag{26}$$

Theorem 5 (Supercritical, $\lambda > 0$) *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda > 0$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x \in E, g \in B_1^+(E)} \left| \varphi(x)^{-1} \mathrm{e}^{-\ell \lambda t} M_t^{(\ell)}[g](x) - \ell! \, \langle g, \tilde{\varphi} \rangle^\ell \, L_\ell \right|,$$

where L_k was defined in Theorem 2, albeit that $L_1 = 1/\lambda$.

Then, for all $\ell \leq k$

$$\sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \to \infty} \Delta_t^{(\ell)} = 0. \tag{27}$$

Theorem 6 (Subcritical, $\lambda < 0$) *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda < 0$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x \in E, g \in B_1^+(E)} \left| \varphi(x)^{-1} M_t^{(\ell)}[g](x) - \ell! \, \langle g, \tilde{\varphi} \rangle^\ell \, L_\ell \right|,$$

where $L_1 = 1/|\lambda|$ and for $k \geq 2$, the constants L_k are defined recursively via

$$L_k = \frac{1}{|\lambda|} \left\langle \beta \mathcal{E} \left[\sum_{[k_1, \ldots, k_N]_k^2} \prod_{\substack{j=1 \\ j:k_j > 0}}^N \varphi(x_j) L_{k_j} \right], \tilde{\varphi} \right\rangle - \frac{\langle g\varphi, \tilde{\varphi} \rangle}{|\lambda| \langle g, \tilde{\varphi} \rangle} L_{k-1},$$

if X is a branching Markov process and

$$L_k(x) = (-1)^k \langle \mathbb{U}_k [\varphi], \tilde{\varphi} \rangle - \mathcal{R}_k(x),$$

where

$$\mathcal{R}_k(x) = \sum_{\{m_1, \ldots, m_{k-1}\}_k} \frac{(-1)^k}{m_1! \ldots m_{k-1}!} (-\varphi(x))^{m_1 + \cdots + m_{k-1} - 1} (m_1 + \cdots + m_{k-1} - 1)!$$

$$\prod_{j=1}^{k-1} L_j(x)^{m_j},$$

and where we define recursively, $\mathbb{U}_1 [\varphi] (x) = -\varphi(x)/|\lambda|$ and for $k \geq 2$,

$$\mathbb{U}_k [\varphi] (x) = \frac{1}{|\lambda|} \left(\mathbb{U}_{k-1} [\varphi] (x) + \sum_{\{m_1, \ldots, m_{k-1}\}_k} \frac{1}{m_1! \ldots m_{k-1}!} \prod_{j=1}^{k-1} \langle \mathbb{U}_j [\varphi], \tilde{\varphi} \rangle^{m_j} \right.$$

$$\times \left[\psi^{(m_1 + \cdots + m_{k-1})}(x, 0+)(-\varphi(x))^{m_1 + \cdots + m_{k-1}} + \beta(x) \right.$$

$$\left. \left. \int_{M(E)^\circ} \langle \varphi, \nu \rangle^{m_1 + \cdots + m_{k-1}} \Gamma(x, d\nu) \right] \right)$$

if X is a superprocess. Then, for all $\ell \leq k$

$$\sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \to \infty} \Delta_t^{(\ell)} = 0. \tag{28}$$

The results in Theorems 4, 5 and 6 are slightly less predictable. Let us discuss this point a little further in the particle setting for convenience. For the supercritical case, the extra "linear" term arising from the time integral does not affect the exponential growth of the process, and hence the leading order behaviour is still dominated by $e^{k\lambda t}$. In the critical case, let us assume momentarily that we are permitted to assume a Yaglom limit holds (see for example [20] in the branching Markov process setting). In that setting, we know that, conditional on $N_t > 0$, $\langle f, X_t \rangle \sim O(t)$ at $t \to \infty$. Conditional on survival, we thus have (up to a constant)

$$\int_0^t \langle f, X_t \rangle \mathrm{d}s \sim \int_0^t s \mathrm{d}s = t^2.$$

This implies that (still conditional on survival) the k-th moment of the occupation measure behaves like t^{2k}. Recalling that, at criticality, the survival probability behaves like $1/t$, we obtain the scaling t^{2k-1}. Finally, in the subcritical case, we know that the total occupation $\int_0^\zeta \langle g, X_s \rangle \mathrm{d}s$, where $\zeta = \inf\{t > 0 : \langle 1, X_t \rangle = 0\}$, is finite, behaving like an average spatial distribution of mass, i.e. $\langle g, \tilde{\varphi} \rangle$, multiplied by ζ, meaning that no normalisation is required to control the "growth" of the running occupation moments in this case.

1.5 Examples for specific branching processes

We now give some examples to illustrate our results and the generality of this setting. **Continuous-time Galton–Watson process and CSBPs:** Let us now consider the simplest branching particle setting where the process is not spatially dependent. In effect, we can take $E = \{0\}$, $\mathbb{P}$ to be the Markov process which remains at $\{0\}$ and the branching mechanism has no spatial dependence. This is the setting of a continuous-time Galton–Watson process. Its branching rate β is constant, and the first and second moments of the offspring distribution are given by $m_1 = \mathcal{E}[N]$ and $m_2 = \mathcal{E}[N^2]$, respectively, where N is the number of offspring produced at a branching event, are key to characterising moment limits. When the process is independent of space, we have $\lambda = \beta(m_1 - 1)$, $\varphi = 1$, $\tilde{\varphi}$ can be taken as $\delta_{\{0\}}$ and (H1) trivially holds. Theorem 1 now tells us that, at criticality, i.e. $m_1 = 1$, the limit for the ℓ-th moment of the population size at time t, i.e. N_t, satisfies

$$t^{-(\ell-1)} \mathbb{E}[N_t^\ell] \sim 2^{-(\ell-1)} \ell! \left(\beta(m_2 - 1)\right)^{\ell-1}, \qquad \text{as } t \to \infty, \tag{29}$$

when $\mathcal{E}[N^\ell] < \infty$ and $\ell \geq 1$.

In the supercritical case, i.e. $m_1 > 1$, the limit in Theorem 2 simplifies to

$$e^{-\beta(m_1-1)\ell t} \mathbb{E}[N_t^\ell] \sim \ell! L_\ell, \qquad \text{as } t \to \infty, \tag{30}$$

where the iteration

$$L_\ell = \frac{1}{(m_1 - 1)(\ell - 1)} \mathcal{E}\left[\sum_{[k_1,\dots,k_N]_\ell^2} \prod_{\substack{j=1 \\ j:k_j>0}}^{N} L_{k_j} \right], \qquad \ell \geq 2,$$

holds. Here, although the simplified formula is still a little complicated, it demonstrates more clearly that the moments in Theorem 2 grow according to the leading order terms of the offspring distribution. Indeed, in the case $\ell = 2$, we have

$$L_2 = \frac{1}{m_1 - 1} \mathcal{E}[\mathrm{card}\{[k_1, \dots, k_N]_2^2\}] = \frac{1}{m_1 - 1} \frac{\mathcal{E}[N(N-1)]}{2} = \frac{m_2 - m_1}{2(m_1 - 1)}.$$

The constant L_3 can now be computed explicitly in terms of L_2 and $L_1 = 1$, and so on.

The limits in the subcritical case can be detailed similarly and only offer minor simplifications of the constants L_k, $k \geq 1$ presented in the statement of Theorem 3. Hence we leave the details for the reader to check.

The analogue of the continuous-time Galton–Watson process in the superprocess setting is that of a continuous-state branching process (CSBP). In this setting there is no associated movement, ψ in (7) is not spatially dependent and the non-local branching mechanism (8) satisfies $\phi \equiv 0$. The right eigenvector and left eigenmeasure can be structured in the same way as in the Galton–Watson setting, with eigenvalue given by $-\psi'(0+) = b$.

Similarly to (29), Theorem 1 tells us at criticality, i.e. $-\psi'(0+) = 0$, that the CSBP $(Z_t, t \geq 0)$ satisfies

$$t^{-(\ell-1)} \mathbb{E}[Z_t^\ell] \sim 2^{-(\ell-1)} \ell! \psi''(0+)^{\ell-1}, \qquad \text{as } t \to \infty,$$

when $\int_0^\infty |y|^\ell \nu(\mathrm{d}y) < \infty$ and $\ell \geq 1$. The situation at super- and subcriticality again do not offer much more insight than the natural analogue of (30) with an induction on the L_k constants. Hence we end our discussion of the CSBP example here.

Crump-Mode-Jagers (CMJ) processes: Consider a branching process in which particles live for a random amount of time ζ and during their lifetime, give birth to a (possibly random) number of offspring at random times; in essence, the age of a parent at the birth times forms a point process, say $\eta(\mathrm{d}t)$ on $[0, \zeta]$. We denote the law of the latter by $\mathcal{P}$. The offspring reproduce and die as independent copies of the parent particle and the law of the process is denotes by $\mathbb{P}$ when initiated from a single individual. Although this model, i.e. a CMJ process, is not covered in the present article, it appears to be the only context in which comparable asymptotic moment results can be found in the literature.

First let us consider the case where the number of offspring, $N = \eta[0, \zeta]$, born to the initial individual during its lifetime satisfies $\mathbb{E}[N] = 1$, which corresponds to the critical case. (Note, criticality is usually described in terms of the Malthusian parameter, however, the setting $\mathbb{E}[N] = 1$ is equivalent.) Further, let Z_t denote the number of individuals in the population at time $t \geq 0$. Under the moment assumption

$\mathcal{E}[N^k] < \infty$ for some $k \geq 1$, [11] showed that the factorial moments $m_k(t) := \mathcal{E}[Z_t(Z_t - 1) \cdots (Z_t - k + 1)]$ satisfy

$$\lim_{t \to \infty} \frac{m_k(t)}{t^{k-1}} = k! \frac{\mathcal{E}[\zeta]^k}{b^{2k-1}} (m_2 - 1)^{k-1},$$

where $m_2 = \mathcal{E}[N^2]$ and $b = \mathcal{E}[\int_0^\zeta t\eta(\mathrm{d}t)]$. We encourage the reader to compare this to the spatially independent example considered above.

The proof in [11] echoes a similar approach to the one presented in the current article. The author first develops a non-linear integral equation that describes the evolution of $m_k(t)$ in terms of the lower order moments, cf. [11,Theorem 1]. An inductive argument along with this evolution equation is then used to prove the above asymptotics.

Branching Brownian motion in a bounded domain: In [28], the Yaglom limit for branching Brownian motion (BBM) in a bounded domain was proved. In this setting, the semigroup P corresponds to that of a Brownian motion killed on exiting a C^1 domain, E. The branching rate is taken as the constant $\beta > 0$ and the offspring distribution is not spatially dependent. Moreover, the first and second moments, $m_1 := \mathcal{E}[N]$ and $m_2 = \mathcal{E}[N^2]$, are finite. In this setting, the right eigenfunction φ exists on E, satisfying Dirichlet boundary conditions, and is accompanied by the left eigenmeasure $\varphi(x)\mathrm{d}x$ on E. The associated eigenvalue is identified explicitly as $\lambda = \beta(m_1 - 1) - \lambda_E$, where λ_E is the ground state eigenvalue of the Laplacian on E. The critical regime thus occurs when $\beta(m_1 - 1) = \lambda_E$

Among the main results of [28], are the Kolmogorov limit,

$$\mathbb{P}_{\delta_x}(N_t > 0) \sim \frac{1}{t} \frac{2(m_1 - 1)\varphi(x)}{\lambda_E(m_2 - m_1) \int_E \varphi(x)^3 \mathrm{d}x} =: 2\varphi(x)/\Sigma, \qquad x \in E,$$

as $t \to \infty$ and the Yaglom distributional limit,

$$\mathrm{Law}\left(\frac{\langle f, X_t \rangle}{t} \,\middle|\, N_t > 0 \right) \to \mathrm{Law}(\mathbf{e}_{2/\langle \varphi, f \rangle \Sigma}), \qquad \text{as } t \to \infty,$$

where $\mathbf{e}_{\langle \varphi, f \rangle \Sigma/2}$ is an exponentially distributed random variable with rate $\langle \varphi, f \rangle \Sigma/2$. (Note, we understand $\langle f, \varphi \rangle = \int_E \varphi(x) f(x) \mathrm{d}x$ in this context.) In particular these two results allude to the limit of moments (albeit further moment assumptions would be needed on N), which, in the spirit of (20), can be heuristically read as

$$\lim_{t \to \infty} \frac{1}{t^{k-1}} \mathbb{E}_{\delta_x}[\langle f, X_t \rangle^k] = \lim_{t \to \infty} t\mathbb{P}_{\delta_x}(N_t > 0)\mathbb{E}_{\delta_x}\left[\frac{\langle f, X_t \rangle^k}{t^k} \,\middle|\, N_t > 0 \right]$$

$$= k!2^{-(k-1)}\langle f, \varphi \rangle^k \Sigma^{k-1} \varphi(x), \qquad x \in E, \qquad (31)$$

for $k \geq 1$. Taking into account the fact that $\tilde{\varphi}(x) = \varphi(x)\mathrm{d}x$ and $\beta = \lambda_E/(m_1 - 1)$, we see that

$$\langle \mathbb{V}[\varphi](x), \tilde{\varphi} \rangle = \langle \beta\varphi^2(m_2 - m_1), \varphi \rangle = \lambda_E \frac{(m_2 - m_1)}{(m_1 - 1)} \int_E \varphi(x)^3 \mathrm{d}x = \Sigma.$$

Hence (31) agrees precisely with Theorem 1.

Neutron branching processes: The neutron branching process (NBP), as introduced in [3, 22], is our final example. Neutrons evolve in the configuration space $E = D \times V$, where $D \subset \mathbb{R}^3$ is a bounded, open set denoting the set of particle locations and $V := \{\upsilon \in \mathbb{R}^3 : \mathrm{v}_{\min} \leq |\upsilon| \leq \mathrm{v}_{\max}\}$ with $0 < \mathrm{v}_{\min} \leq \mathrm{v}_{\max} < \infty$, denotes the set of velocities. From an initial space-velocity configuration (r, υ), particles move according to piecewise deterministic Markov processes characterised by $\sigma_\mathrm{s}\pi_\mathrm{s}$, where $\sigma_\mathrm{s}(r, \upsilon)$ denotes the rate at which particles change velocity (also called *scattering* events) at (r, υ), and $\pi_\mathrm{s}(r, \upsilon, \upsilon')\mathrm{d}\upsilon'$ denotes the probability that such a scattering event results in a new outgoing velocity υ'. When at $(r, \upsilon) \in D \times V$, at rate $\sigma_\mathrm{f}(r, \upsilon)$, a branching (or *fission*) event occurs, resulting in the release of several new neutrons with configurations $(r, \upsilon_1), \ldots, (r, \upsilon_N)$, say. The quantity $\pi_\mathrm{f}(r, \upsilon, \upsilon')\mathrm{d}\upsilon'$ gives the average number of neutrons produced with outgoing velocity υ' from a fission event at (r, υ). Thus, the NBP is an example of a branching Markov process with non-local branching, where the motion is a piecewise deterministic Markov process and the non-locality at branching events appears in the velocity.

As previously mentioned, Theorem 1 was established for the NBP in [20] using a spine decomposition approach but under more restrictive assumptions and only when the test function f is taken to be the right eigenfunction φ. In [22], under the assumptions

(A1) $\sigma_\mathrm{s}, \pi_\mathrm{s}, \sigma_\mathrm{f}, \pi_\mathrm{f}$ are uniformly bounded from above.
(A2) $\inf_{r \in D\upsilon, \upsilon' \in V}(\sigma_\mathrm{s}(r, \upsilon)\pi_\mathrm{s}(r, \upsilon, \upsilon') + \sigma_\mathrm{f}(r, \upsilon)\pi_\mathrm{f}(r, \upsilon, \upsilon')) > 0,$

it was shown that (H1) holds. Moreover, since only a finite number of neutrons can be produced at a fission event, the number of offspring is uniformly bounded from above and thus (H2) holds for all $k \geq 1$. Hence, the results obtained in this paper hold for the NBP.

Although there is no particular simplification of the limiting constants in Theorems 1, 2, 3, 4, 5 and 6 in this setting, of particular interest is the notion of particle clustering that appears in Monte Carlo criticality calculations [8, 9, 33]. This phenomena occurs in critical reactors where particles exhibit strong spatial correlations. Studying the moments $\mathrm{M}^{(k)}$ and other correlation structures in this setting will shed further light on this phenomenon.

1.6 Outline of the paper and strategy for proofs

In the remainder of the paper, there are two sections. One handling the proofs for the branching Markov process setting, and another handling the proofs for the superprocess setting. The proofs of the six main theorems follow a similar pattern and so we will briefly provide a heuristic to their strategic approach here, and how this is laid

out in the remainder of the paper. Roughly speaking, each of the proofs follows the following fundamental steps.

Step 1: The first step in all of the proofs is to establish a non-linear semigroup equation in the spirit of (4) for branching Markov processes and (5) for superprocesses, but for the joint Markovian pair $\langle f, X_t \rangle$ and $\int_0^t \langle g, X_s \rangle \mathrm{d}s$, $t \geq 0$. Moreover, we want this semigroup equation to be written as an integral equation in terms of $(\mathrm{T}_t, t \geq 0)$, rather than $(\mathrm{P}_t, t \geq 0)$; recall $\mathrm{T} = \mathrm{T}^{(1)}$ is the mean semigroup cf. (13). This is done for branching Markov processes in detail in Sect. 2.1, whereas we simply state the relationship for superprocesses in Sect. 3.1 as the proof is very similar.

Step 2: Next, we work with the simple observation that differentiating the non-linear semigroup k times will give us access to the the k-th moment process $(\mathrm{T}_t^{(k)}, t \geq 0)$. That is,

$$\mathrm{T}_t^{(k)}[f](x) = (-1)^k \frac{\partial^k}{\partial \theta^k} \mathbb{E}_{\delta_x}[\mathrm{e}^{-\theta \langle f, X_t \rangle}]\Big|_{\theta=0},$$

with a similar formula holding for $\mathrm{M}^{(k)}$. However, given the recursion of the non-linear semigroup derived in the previous step, this turns out to give us a new recursion of $\mathrm{T}^{(k)}$ in terms of $\mathrm{T}^{(k-1)}, \mathrm{T}^{(k-2)}, \ldots, \mathrm{T}^{(1)}$ and similarly for $\mathrm{M}^{(k)}$ in terms of $\mathrm{M}^{(k-1)}, \mathrm{M}^{(k-2)}, \ldots, \mathrm{M}^{(1)}$; the two recursions are extremely close to one another, but subtly different nonetheless.

Here we see a fundamental difference between the approaches we take for superprocesses and branching Markov processes. This is due to the way in which we differentiate the branching mechanisms in the recursion of the non-linear semigroup derived in Step 1. For the branching Markov process setting, the branching mechanism G that appears in the aforesaid recursion is written in terms of a product over particles at a birth event. This allows us to use the Leibniz rule when differentiating across the product; this is done in Sect. 2.2. In the case of superprocesses, the branching mechanism is written in terms of an analytically smoother object, namely a Lévy-Khintchine-type formula. With no representation in terms of particles, we must turn instead to Fàa di Bruno's rule in order to differentiate this branching mechanism; cf. Sect. 3.2

Step 3: The differentiation in Step 2 yields two fundamentally different sets of equations for the k-th moment evolutions and k-th occupation moment evolutions for each of the two classes of branching processes. In turn, this results in slightly different combinatorial formulae to work with when when completing the proofs of the theorems. In essence the proof of all of the main results now pertains to identifying the leading order terms in the aforesaid combinatorial formulae and appealing to an inductive argument; for example, using the statement of Theorem 1 for $\ell \leq k$ to prove the statement of that theorem for $\ell = k + 1$.

In Sect. 2.3 we give the full argument for the proof of Theorem 1 in the branching Markov process setting. This sets the scene for all the other proofs for this class of branching processes. Indeed, based on this proof, we then complete the proof of Theorem 4 in Sect. 2.4, and Theorems 2, 3, 5 and 6 in Sect. 2.5, highlighting only the main differences relative to the key arguments in the proof of Theorem 1.

In the same spirit, we give a complete proof for Theorem 1 in Sect. 3.3 for super-processes which gives the roadmap for all other proofs for this class of processes. Thereafter, in Sect. 3.4, we highlight only the differences for the proofs of Theorems 2 and 3 for superprocesses. Finally, at even greater brevity than in the branching Markov process setting (because of familiarity) we sketch the main differences for the proof of Theorem 4, 5, and 6.

2 Proofs for branching Markov processes

The proof is a mixture of analytical and combinatorial computations which are based around the behaviour of the linear and non-linear semigroups of X.

2.1 Linear and non-linear semigroup equations

For $f \in B^+(E)$, it is well known that the mean semigroup evolution satisfies

$$\mathrm{T}_t[f](x) = \mathrm{P}_t[f](x) + \int_0^t \mathrm{P}_s \left[\mathrm{F}\mathrm{T}_{t-s}[f] \right](x)\mathrm{d}s \qquad t \geq 0, x \in E, \qquad (32)$$

where

$$\mathrm{F}[f](x) = \beta(x)\mathcal{E}_x \left[\sum_{i=1}^N f(x_i) - f(x) \right] =: \beta(x)(\mathrm{m}[f](x) - f(x)), \qquad x \in E.$$

See for example the calculations in [22]. Associated with every linear semigroup of a branching process is a so-called many-to-one formula. Many-to-one formulae are not necessarily unique and the one we will develop here is slightly different from the usual construction because of non-locality.

Suppose that $\xi = (\xi_t, t \geq 0)$, with probabilities $\mathbf{P} = (\mathbf{P}_x, x \in E)$, is the Markov process corresponding to the semigroup P. Let us introduce a new Markov process $\hat{\xi} = (\hat{\xi}_t, t \geq 0)$ which evolves as the process ξ but at rate $\beta(x)\mathrm{m}[1](x)$ the process is sent to a new position in E, such that for all Borel $A \subset E$, the new position is in A with probability $\mathrm{m}[\mathbf{1}_A](x)/\mathrm{m}[1](x)$. We will refer to the latter as *extra jumps*. Note the law of the extra jumps is well defined thanks to the assumption (15), which ensures that $\sup_{x \in E} \mathrm{m}[1](x) = \sup_{x \in E} \mathcal{E}_x(\langle 1, \mathcal{Z} \rangle) < \infty$. Accordingly we denote the probabilities of $\hat{\xi}$ by $(\hat{\mathbf{P}}_x, x \in E)$. We can now state our many-to-one formula.

Lemma 1 *Write $B(x) = \beta(x)(\mathrm{m}[1](x) - 1)$, $x \in E$. For $f \in B^+(E)$ and $t \geq 0$, we have*

$$\mathrm{T}_t[f](x) = \hat{\mathbf{E}}_x \left[\exp \left(\int_0^t B(\hat{\xi}_s)\mathrm{d}s \right) f(\hat{\xi}_t) \right]. \qquad (33)$$

The proof is classical and follows standard reasoning for semigroup integral equations e.g. as in [19, 22]: First conditioning the right-hand side of (33) on the time of

the first extra jump, then using the principle of transferring between multiplicative and additive potentials in the resulting integral equation (cf. Lemma 1.2, Chapter 4 in [13]) shows that (32) holds. Grönwall's Lemma, the fact that $\beta \in B^+(E)$ and (15) for $k = 1$ ensure that the relevant integral equations have unique solutions.

We now define a variant of the non-linear evolution equation (2) associated with X via

$$\mathrm{u}_t[f, g](x) = \mathbb{E}_{\delta_x}\left[1 - \mathrm{e}^{-\langle f, X_t\rangle - \int_0^t \langle g, X_s\rangle \mathrm{d}s}\right], \qquad t \geq 0, \ x \in E, \ f, g \in B^+(E). \quad (34)$$

For $f \in B_1^+(E)$, define

$$\mathrm{A}[f](x) = \beta(x)\mathcal{E}_x\left[\prod_{i=1}^{N}(1 - f(x_i)) - 1 + \sum_{i=1}^{N} f(x_i)\right], \qquad x \in E.$$

Our first preparatory result relates the two semigroups $(\mathrm{u}_t, t \geq 0)$ and $(\mathrm{T}_t, t \geq 0)$.

Lemma 2 *For all $f, g \in B^+(E)$, $x \in E$ and $t \geq 0$, the non-linear semigroup $\mathrm{u}_t[f, g](x)$ satisfies*

$$\mathrm{u}_t[f, g](x) = \mathrm{T}_t[1 - \mathrm{e}^{-f}](x) - \int_0^t \mathrm{T}_s\left[\mathrm{A}[\mathrm{u}_{t-s}[f, g]] - g(1 - \mathrm{u}_{t-s}[f, g])\right](x)\mathrm{d}s. \quad (35)$$

Proof Again, the proof uses standard techniques for integral evolution equations so we only sketch the proof. Instead of considering $\mathrm{u}_t[f, g]$, we will first work instead with

$$\mathrm{v}_t[f, g] = \mathbb{E}_{\delta_x}\left[\mathrm{e}^{-\langle f, X_t\rangle - \int_0^t \langle g, X_s\rangle \mathrm{d}s}\right], \qquad t \geq 0, \ x \in E, \ f, g \in B^+(E), \quad (36)$$

which will turn out to be more convenient for technical reasons.

By splitting the expectation in (36) on the first branching event and appealing to the Markov property, we get, for $f, g \in B^+(E)$, $t \geq 0$ and $x \in E$,

$$\mathrm{v}_t[f, g](x) = \mathbf{E}_x\left[\mathrm{e}^{-\int_0^t \beta(\xi_s)\mathrm{d}s}\mathrm{e}^{-f(\xi_t) - \int_0^t g(\xi_s)\mathrm{d}s}\right]$$
$$+ \mathbf{E}_x\left[\int_0^t \beta(\xi_s)\mathrm{e}^{-\int_0^s \beta(\xi_u) + g(\xi_u)\mathrm{d}u}\mathrm{H}[\mathrm{v}_{t-s}[f, g]](\xi_s)\mathrm{d}s\right],$$

where

$$\mathrm{H}[g](x) = \mathcal{E}_x\left[\prod_{i=1}^{N} g(x_i)\right], \qquad g \in B^+(E), x \in E.$$

Using similar reasoning to Lemma 1.2, Chapter 4 in [13] we can move the multiplicative potential with rate $\beta + g$ to an additive potential in the above evolution equation

to obtain

$$\mathrm{v}_t[f, g](x) = \hat{\mathrm{P}}_t[\mathrm{e}^{-f}](x) + \int_0^t \mathrm{P}_s \left[\mathrm{G}[\mathrm{v}_{t-s}[f, g]] - g\mathrm{v}_{t-s}[f, g] \right](x)\mathrm{d}s. \quad (37)$$

Now define

$$\mathrm{D}[f](x) = \beta(x)\mathcal{E}_x \left[\prod_{i=1}^N f(x_i) - \sum_{i=1}^N f(x_i) \right]$$
$$= \beta(x) \left(\mathrm{H}[f](x) - \mathrm{m}[f](x) \right), \quad f \in B_1^+(E), x \in E$$

and $(\tilde{\mathrm{v}}_t, t \geq 0)$ via

$$\tilde{\mathrm{v}}_t[f, g](x) = \mathrm{T}_t[\mathrm{e}^{-f}](x) + \int_0^t \mathrm{T}_s \left[\mathrm{D}[\tilde{\mathrm{v}}_{t-s}[f, g]] - g\tilde{\mathrm{v}}_{t-s}[f, g] \right](x)\mathrm{d}s$$
$$= \hat{\mathbf{E}}_x \left[\mathrm{e}^{\int_0^t \mathrm{B}(\hat{\xi}_s)\mathrm{d}s} \mathrm{e}^{-f(\hat{\xi}_t)} \right]$$
$$+ \hat{\mathbf{E}}_x \left[\int_0^t \mathrm{e}^{\int_0^s \mathrm{B}(\hat{\xi}_u)\mathrm{d}u} \left(\mathrm{D}[\tilde{\mathrm{v}}_{t-s}[f, g]](\hat{\xi}_s) - g(\hat{\xi}_s)\tilde{\mathrm{v}}_{t-s}[f, g](\hat{\xi}_s) \right) \mathrm{d}s \right],$$
(38)

for $x \in E, t \geq 0$ and $f, g \in B^+(E)$. Note that for the moment we don't claim a solution to (38) exists.

For convenience, we will define

$$\mathrm{K}_t[f, g](x) = \hat{\mathbf{E}}_x \left[\int_0^t \mathrm{e}^{\int_0^s \mathrm{B}(\hat{\xi}_u)\mathrm{d}u} \left(\mathrm{D}[\tilde{\mathrm{v}}_{t-s}[f, g]](\hat{\xi}_s) - g(\hat{\xi}_s)\tilde{\mathrm{v}}_{t-s}[f, g](\hat{\xi}_s) \right) \mathrm{d}s \right],$$

so that $\tilde{\mathrm{v}}_t[f, g](x) = \mathrm{T}_t[\mathrm{e}^{-f}](x) + \mathrm{K}_t[f, g](x)$. By conditioning the right-hand side of (38) on the first jump of $\hat{\xi}$ (bearing in mind the dynamics of $\hat{\xi}$ given just before Lemma 1) with the help of the Markov property (recalling that $\mathrm{B}(x) - \beta\mathrm{m}[1] = \beta$), we get

$$\tilde{\mathrm{v}}_t[f, g](x)$$
$$= \mathbf{E}_x \left[\mathrm{e}^{-\int_0^t \beta(\xi_s)\mathrm{d}s} \mathrm{e}^{-f(\xi_t)} \right] + \mathbf{E}_x \left[\int_0^t \beta(\xi_\ell)\mathrm{m}[1](\xi_\ell)\mathrm{e}^{-\int_0^\ell \beta(\xi_s)\mathrm{d}s} \frac{\mathrm{m}[\mathrm{T}_{t-\ell}[\mathrm{e}^{-f}]](\xi_\ell)}{\mathrm{m}[1](\xi_\ell)}\mathrm{d}\ell \right]$$
$$+ \mathbf{E}_x \left[\mathrm{e}^{-\int_0^t \beta(\xi_u)\mathrm{m}[1](\xi_u)\mathrm{d}u} \int_0^t \mathrm{e}^{\int_0^s \mathrm{B}(\xi_u)\mathrm{d}u} \left(\mathrm{D}[\tilde{\mathrm{v}}_{t-s}[f, g]](\xi_s) - g(\xi_s)[\tilde{\mathrm{v}}_{t-s}[f, g]) \mathrm{d}s \right] \right.$$
$$+ \mathbf{E}_x \left[\int_0^t \beta(\xi_\ell)\mathrm{m}[1](\xi_\ell)\mathrm{e}^{-\int_0^\ell \beta(\xi_u)\mathrm{m}[1](\xi_u)\mathrm{d}u} \right.$$

$$\left(\int_0^\ell \mathrm{e}^{\int_0^s \mathrm{B}(\xi_u)\mathrm{d}u} \left(\mathrm{D}[\tilde{\mathrm{v}}_{t-s}[f, g]](\xi_s) - g(\xi_s)\mathrm{v}_{t-s}[f, g](\xi_s) \right) \mathrm{d}s \right.$$
$$\left. + \mathrm{e}^{\int_0^\ell \mathrm{B}(\xi_u)\mathrm{d}u} \frac{\mathrm{m}[\mathrm{K}_{t-\ell}[g]](\xi_\ell)}{\mathrm{m}[1](\xi_\ell)} \right) \mathrm{d}\ell \right].$$

Gathering terms and exchanging the order of integration in the double integral, this simplifies to

$$\tilde{v}_t[f, g](x)$$

$$= \mathbf{E}_x \left[e^{-\int_0^t \beta(\xi_s)ds} e^{-f(\xi_t)} \right] + \mathbf{E}_x \left[\int_0^t \beta(\xi_\ell) e^{-\int_0^\ell \beta(\xi_s)ds} \mathrm{m}[\tilde{v}_{t-\ell}[f, g](x)](\xi_\ell) d\ell \right]$$

$$\mathbf{E}_x \left[e^{-\int_0^t \beta(\xi_u)\mathrm{m}[1](\xi_u)du} \int_0^t e^{\int_0^s \mathrm{B}(\xi_u)du} \left(\mathrm{D}[\tilde{v}_{t-s}[f, g]](\xi_s) - g(\xi_s)[\tilde{v}_{t-s}[f, g]] \right) ds \right]$$

$$+ \mathbf{E}_x \left[\int_0^t \int_0^t \mathbf{1}_{(s \le \ell)} \beta(\xi_\ell) \mathrm{m}[1](\xi_\ell) e^{-\int_0^\ell \beta(\xi_u)\mathrm{m}[1](\xi_u)du} e^{\int_0^s \mathrm{B}(\xi_u)du} \right.$$

$$\left. \left(\mathrm{D}[\tilde{v}_{t-s}[f, g]](\xi_s) - g(\xi_s)\tilde{v}_{t-s}[f, g](\xi_s)) \right) d\ell \, ds \right]$$

$$= \mathbf{E}_x \left[e^{-\int_0^t \beta(\xi_s)ds} e^{-f(\xi_t)} \right] + \mathbf{E}_x \left[\int_0^t \beta(\xi_\ell) e^{-\int_0^\ell \beta(\xi_s)ds} \mathrm{m}[\tilde{v}_{t-\ell}[g](x)](\xi_\ell) d\ell \right]$$

$$+ \mathbf{E}_x \left[\int_0^t e^{-\int_0^s \beta(\xi_u)du} \left(\mathrm{D}[\tilde{v}_{t-s}[f, g]](\xi_s) - g(\xi_s)\tilde{v}_{t-s}[f, g](\xi_s)) \right) ds \right].$$

Finally, appealing to the change of multiplicative potential to additive potential in the spirit of e.g. Lemma 1.2, Chapter 4 of [13], we get

$$\tilde{v}_t[f, g](x) = \hat{\mathrm{P}}_t[e^{-f}](x) + \int_0^t \mathrm{P}_t \left[\mathrm{G}[\tilde{v}_{t-s}[f, g]] - g\tilde{v}_{t-s}[f, g] \right](x) ds$$

and hence $(\tilde{v}_t, t \ge 0)$ is a solution to (37). Reversing these arguments also shows that solutions to (37) solve (38). As such, a standard argument using $\beta \in B^+(E)$, the assumption (15) for $k = 1$ and Grönwall's Lemma tells us that all of the integral equations thus far have unique solutions. In conclusion, $(v_t[g], t \ge 0)$ and $(\tilde{v}_t[g], t \ge 0)$ agree.

To complete the lemma, note that

$$1 - \mathrm{T}_t[e^{-f}](x) = \mathrm{T}_t[1 - e^{-f}](x) + 1 - \mathrm{T}_t[1](x)$$

moreover,

$$1 - \mathrm{T}_t[1](x) = \hat{\mathbf{E}}_x \left[\int_0^t \mathrm{B}(\hat{\xi}_s) e^{\int_0^s \mathrm{B}(\hat{\xi}_u)du} ds \right] = \int_0^t \mathrm{T}_s[\mathrm{B}](x) ds.$$

Hence, working form (38) and the definitions of D and A, which are related via

$$\mathrm{D}[1 - f](x) = \beta(x)\mathcal{E}_x \left[\prod_i (1 - f(x_i)) - \sum_{i=1}^N (1 - f(x_i)) \right]$$

$$= \mathrm{A}[f](x) + \mathrm{B}(x), \qquad x \in E, f \in B_1^+(E),$$

we get

$$
\begin{aligned}
u_t[f, g](x) &= 1 - v_t[f, g](x) \\
&= 1 - T_t[e^{-f}](x) \\
&\quad - \int_0^t T_s \left[D[1 - u_{t-s}[f, g]] - g(1 - u_{t-s}[f, g]) \right](x) ds \\
&= T_t[1 - e^{-f}](x) - \int_0^t T_s \left[A[u_{t-s}[f, g]] - g(1 - u_{t-s}[f, g]) \right](x) ds,
\end{aligned}
$$

as required. $\qquad \square$

2.2 Evolution equations for the *k*-th moment of branching Markov processes

Next we turn our attention to the evolution equation generated by the k-th moment functional $T_t^{(k)}$, $t \geq 0$. To this end, we start by observing that

$$
T_t^{(k)}[f](x) = (-1)^{k+1} \frac{\partial^k}{\partial \theta^k} u_t[e^{-\theta f}](x) \Big|_{\theta=0}. \tag{39}
$$

The following result gives us an iterative approach to writing the k-th moment functional in terms of lower order moment functionals.

Proposition 1 *Fix $k \geq 2$. Under the assumptions of Theorem 1, with the additional assumption that*

$$
\sup_{x \in E, s \leq t} T_s^{(\ell)}[f](x) < \infty, \qquad \ell \leq k - 1, f \in B^+(E), t \geq 0, \tag{40}
$$

it holds that

$$
T_t^{(k)}[f](x) = T_t[f^k](x) + \int_0^t T_s \left[\beta \eta_{t-s}^{(k-1)}[f] \right](x) \, ds, \qquad t \geq 0, \tag{41}
$$

where

$$
\eta_{t-s}^{(k-1)}[f](x) = \mathcal{E}_x \left[\sum_{[k_1, \ldots, k_N]_k^2} \binom{k}{k_1, \ldots, k_N} \prod_{j=1}^N T_{t-s}^{(k_j)}[f](x_j) \right],
$$

and $[k_1, \ldots, k_N]_k^2$ is the set of all non-negative N-tuples $(k_1, \ldots, k_N)$ such that $\sum_{i=1}^N k_i = k$ and at least two of the k_i are strictly positive.

Proof Recall from (35) that

$$
u_t[\theta f, 0](x) = T_t[1 - e^{-\theta f}](x) - \int_0^t T_s \left[A[u_{t-s}[\theta f, 0]] \right](x) ds, \qquad t \geq 0. \tag{42}
$$

It is clear that differentiating the first term k times and setting $\theta = 0$ on the right-hand side of (42) yields

$$\frac{\partial^k}{\partial \theta^k} T_t [1 - e^{-\theta f}](x) \bigg|_{\theta=0} = (-1)^{k+1} T[f^k](x). \tag{43}$$

Thus it remains to differentiate the second term on the right-hand side of (42) k times. To this end, without concern for passing derivatives through expectations, using the Leibniz rule in Lemma A.2 of the Appendix, we have

$$-\frac{\partial^k}{\partial \theta^k} A[u_t[\theta f, 0]](x) \bigg|_{\theta=0}$$

$$= \frac{\partial^k}{\partial \theta^k} \beta(x) \mathcal{E}_x \left[1 - \prod_{i=1}^{N} \mathbb{E}_{\delta_{x_i}} [e^{-\theta \langle f, X_t \rangle}] - \sum_{i=1}^{N} \mathbb{E}_{\delta_{x_i}} [1 - e^{-\theta \langle f, X_t \rangle}] \right]$$

$$= -\beta(x) \mathcal{E}_x \left[\sum_{k_1 + \cdots + k_N = k} \binom{k}{k_1, \ldots, k_m} \prod_{j=1}^{N} (-1)^{k_j} T_t^{(k_j)} [f](x_j) \right.$$

$$\left. + (-1)^{k+1} \sum_{i=1}^{N} T_t^{(k)} [f](x_i) \right]$$

$$= \beta(x) \mathcal{E}_x \left[(-1)^{k+1} \sum_{k_1 + \cdots + k_N = k} \binom{k}{k_1, \ldots, k_m} \prod_{j=1}^{N} T_t^{(k_j)} [f](x_j) \right.$$

$$\left. + (-1)^k \sum_{i=1}^{N} T_t^{(k)} [f](x_i) \right]. \tag{44}$$

where the sum is taken over all non-negative integers $k_1, \ldots, k_N$ such that $\sum_{i=1}^{N} k_i = k$.

Next let us look in more detail at the sum/product term on the right-hand (44). Consider the terms where only one of the k_i in the sum is positive, in which case $k_i = k$ and

$$\binom{k}{k_1, \ldots, k_m} = 1.$$

There are N ways this can happen in the sum of the sum-product term and hence

$$\sum_{k_1 + \cdots + k_N = k} \binom{k}{k_1, \ldots, k_m} \prod_{j=1}^{N} T_t^{(k_j)} [f](x_j)$$

$$= \sum_{i=1}^{N} T^{(k)} [f](x_i) + \sum_{[k_1, \ldots, k_N]_k^2} \binom{k}{k_1, \ldots, k_N} \prod_{j=1}^{N} T_t^{(k_j)} [f](x_j),$$

where $[k_1, \ldots, k_N]_k^2$ is the set of all non-negative N-tuples $(k_1, \ldots, k_N)$ such that $\sum_{i=1}^{N} k_i = k$ and at least two of the k_i are strictly positive. Substituting this back into (44) yields

$$
-\frac{\partial^k}{\partial \theta^k} A[u_t[e^{-\theta f}]]\Big|_{\theta=0}
$$
$$
= (-1)^{k+1} \beta(x) \mathcal{E}_x \left[\sum_{[k_1,\ldots,k_N]_k^2} \binom{k}{k_1, \ldots, k_N} \prod_{j=1}^{N} \mathrm{T}_t^{(k_j)}[f](x_j) \right].
$$

Now let us return to the justification that we can pass the derivatives through the expectation in the above calculation, we first note that derivatives are limits and so an 'epsilon-delta' argument will ultimately require dominated convergence. This is where the assumption (15) and (40) come in. On the right-hand side of (44), each of the $\mathrm{T}_t^{(k_j)}[f](x_j)$ in the sum term are uniformly bounded by the assumption (40) and the collection $[k_1, \ldots, k_N]_k^2$ means that $0 \le k_j \le k - 1$ for each $j = 1, \ldots, N$. Moreover, there can be at most k items in the sum/product. Noting that

$$
\sum_{k_1 + \cdots + k_N = k} \binom{k}{k_1, \ldots, k_m} = N^k, \tag{45}
$$

the assumption (15) allows us to use a domination argument with the k-th order moment.

Combining this with (43) and (42), using an easy dominated convergence argument to pull the k derivatives through the integral in t, then dividing by $(-1)^{k+1}$, we get (41), as required. $\qquad \square$

2.3 Completing the proof of Theorem 1: critical case

We will prove Theorem 1 by induction, starting with the case $k = 1$. In this case, (18) reads

$$
\sup_{t \ge 0} \Delta_t < \infty \quad \text{and} \quad \lim_{t \to \infty} \Delta_t = 0,
$$

which holds due to (14).

We now assume that the theorem holds true in the branching Markov process setting for some $k \ge 1$ and proceed to show that (18) holds for all $\ell \le k + 1$. As alluded to in the introduction, the strategy for the proof is to use equation (40), since this allows us to write the k-th moment in terms of the lower order moments. To prove that the right-hand side of this equation converges to the limit appearing in the statement of the theorem, we use Theorem A.1 stated in the appendix along with the induction hypothesis.

To this end, first note that the induction hypothesis implies that (40) holds. Hence Proposition 1 tells us that

$$\varphi(x)^{-1}t^{-k}\mathrm{T}_t^{(k+1)}[f](x)$$
$$= \varphi(x)^{-1}t^{-k}\mathrm{T}_t[f^{(k+1)}](x)$$
$$+ \varphi(x)^{-1}t^{-k}\int_0^t \mathrm{T}_s\left[\mathcal{E}.\left[\sum_{[k_1,\ldots,k_N]_{k+1}^2}\binom{k+1}{k_1,\ldots,k_N}\prod_{j=1}^N \mathrm{T}_{t-s}^{(k_j)}[f](x_j)\right]\right](x)\mathrm{d}s$$
$$= \varphi(x)^{-1}t^{-k}\mathrm{T}_t[f^{(k+1)}](x) + \varphi(x)^{-1}t^{-(k-1)}$$
$$\int_0^1 \mathrm{T}_{ut}\left[\mathcal{E}.\left[\sum_{[k_1,\ldots,k_N]_{k+1}^2}\binom{k+1}{k_1,\ldots,k_N}\prod_{j=1}^N \mathrm{T}_{t(1-u)}^{(k_j)}[f](x_j)\right]\right](x)\mathrm{d}u, \qquad (46)$$

where we have used the change of variables $s = ut$ in the final equality.

We now make some observations that will simplify the expression on the right-hand side of (46) as $t \to \infty$. First note that due to (14), the first term on the right-hand side of (46) will vanish as $t \to \infty$. Next, note that, if more than two of the k_i in the sum are strictly positive, then the renormalising by t^{k-1} will cause the associated summand to go to zero as well. For example, suppose without loss of generality that k_1 and k_2 are both strictly positive, we can write $t^{k-1} = t^{(k+1)-2} = t^{k_1-1}t^{k_2-1}t^{k_3}\ldots t^{k_N}$. Now the induction hypothesis tells us that the correct normalisation of each of the terms in the product is t^{k_j-1}, which means that the item $\mathrm{T}_{t(1-u)}^{(k_j)}$ for a third $k_j > 0$ will be 'over normalised to zero' in the limit.

To make this heuristic rigorous, we can employ Theorem A.1 from the Appendix. To this end, let us set

$$F[f](x, u, t) := \frac{1}{\varphi(x)t^{k-1}}\mathcal{E}_x\left[\sum_{[k_1,\ldots,k_N]_{k+1}^3}\binom{k+1}{k_1,\ldots,k_N}\prod_{j=1}^N \mathrm{T}_{t(1-u)}^{(k_j)}[f](x_j)\right]$$
$$(47)$$

where $[k_1,\ldots,k_N]_{k+1}^3$ is the subset of $[k_1,\ldots,k_N]_{k+1}^2$, for which at least three of the k_i are strictly positive (which can be an empty set). We will show that conditions (A.1) and (A.2) are satisfied via

$$\sup_{x\in E, f\in B_1^+(E), u\in[0,1]} \varphi(x)F[f](x, u, t) < \infty \text{ and}$$
$$\lim_{t\to\infty} \sup_{u\in[0,1], f\in B_1^+(E), x\in E} \varphi(x)F[f](x, u, t) = 0. \qquad (48)$$

First note that there are no more than $k + 1$ of the k_i that are strictly greater than 1 in the product in (47). This follows from the fact that it is not possible to partition the set

$\{1, \ldots, k+1\}$ into more than $k+1$ non-empty blocks. Next note that

$$\frac{1}{t^{k-1}} \prod_{\substack{j=1 \\ j:k_j>0}}^{N} \mathrm{T}_{t(1-u)}^{(k_j)}[f](x_j) = \frac{(t(1-u))^{k+1-\#\{j:k_j>0\}}}{t^{k-1}}$$

$$\prod_{\substack{j=1 \\ j:k_j>0}}^{N} \varphi(x_j) \cdot \frac{1}{\varphi(x_j)} \frac{\mathrm{T}_{t(1-u)}^{(k_j)}[f](x_j)}{(t(1-u))^{k_j-1}}.$$

The product term on the right-hand side is uniformly bounded in x_j and $t(1-u)$ on compact intervals due to boundedness of φ and the fact that (18) is assumed to hold for all $\ell \leq k$ by induction. Moreover, if $\#\{j : k_j > 0\} \leq 1$, the set $[k_1, \ldots, k_N]_{k+1}^3$ is empty, otherwise, the term $(t(1-u))^{k+1-\#\{j:k_j>0\}}/t^{k-1}$ is finite for all $t \geq 1$, say. From (45) and (15), we also observe that

$$\sup_{x \in E} \mathcal{E}_x \left[\sum_{[k_1,\ldots,k_N]_{k+1}^3} \binom{k+1}{k_1, \ldots, k_N} \right] \leq \sup_{x \in E} \mathcal{E}_x \left[\langle 1, \mathcal{Z} \rangle^{k+1} \right] < \infty.$$

Taking these facts into account, it is now straightforward to see that the earlier given heuristic can be made rigorous and (48) holds. In particular, we can use dominated convergence to pass the limit in t through the expectation in (47) to achieve the second statement in (48).

As F belongs to the class of functions $\mathcal{C}$, defined just before Theorem A.1 in the Appendix, the aforesaid theorem tells us that

$$\lim_{t \to \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \int_0^1 \mathrm{T}_{ut}[\varphi F[f](\cdot, u, t)](x)\mathrm{d}u \right| = 0. \tag{49}$$

Returning to (46), since the sum there requires that at least two of the k_i are positive, this means that the only surviving terms in the limit are those that are combinations of two strictly positive terms k_i and k_j such that $i \neq j$ and $k_i + k_j = k + 1$. This can be thought of as choosing $i, j \in \{1, \ldots N\}$ with $i \neq j$, choosing $k_i \in \{1, \ldots, k\}$ and then setting $k_j = k + 1 - k_i$. One should take care however to avoid double counting each pair (k_i, k_j). Thus, we have

$$\frac{1}{t^k \varphi(x)} \mathrm{T}_t^{(k+1)}[f](x) = \frac{1}{\varphi(x)} \int_0^1 \mathrm{T}_{ut} \left[\frac{\beta(\cdot)}{2t^{k-1}} \mathcal{E}. \left[\sum_{\substack{i=1 \\ j \neq i}}^{N} \sum_{j=1}^{N} \sum_{k_i=1}^{k} \binom{k+1}{k_i, \, k+1-k_i} \right. \right.$$

$$\left. \left. \times \mathrm{T}_{t(1-u)}^{(k_i)}[f](x_i) \mathrm{T}_{t(1-u)}^{(k+1-k_i)}[f](x_j) \right] \right] (x)\mathrm{d}u, \tag{50}$$

where the factor of $1/2$ appears to compensate for the aforementioned double counting.

In order to show that the right-hand side above delivers the required finiteness and limit (18), we again turn to Theorem A.1. For $x \in E$, $t \geq 0$ and $0 \leq u \leq 1$, in anticipation of using this theorem, we now re-define

$$F[f](x, u, t) := \frac{\beta(x)}{2\varphi(x)t^{k-1}}$$
$$\mathcal{E}_x\left[\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \sum_{k_i=1}^{k} \binom{k+1}{k_i, \, k+1-k_i} T_{t(1-u)}^{(k_i)}[f](x_i) T_{t(1-u)}^{(k+1-k_i)}[f](x_j) \right].$$

After some rearrangement, we have

$$F[f](x, u, t)$$
$$= \frac{\beta(x)(1-u)^{k-1}}{2\varphi(x)} \mathcal{E}_x\left[\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \sum_{k_i=1}^{k} \binom{k+1}{k_i, \, k+1-k_i} \right.$$
$$\left. \times \varphi(x_i)\varphi(x_j) \frac{T_{t(1-u)}^{(k_i)}[f](x_i)}{\varphi(x_i)(t(1-u))^{k_i-1}} \frac{T_{t(1-u)}^{(k+1-k_i)}[f](x_j)}{\varphi(x_j)(t(1-u))^{k-k_i}} \right]. \quad (51)$$

Using similar arguments to those given previously in the proof of (49) may, again, combine the induction hypothesis, simple combinatorics and dominated convergence to pass the limit as $t \to \infty$ through the expectation and show that

$$F[f](x, u) := \lim_{t\to\infty} F[f](x, u, t)$$
$$= (k+1)!(\langle \tilde{\varphi}, \mathbb{V}[\varphi] \rangle / 2)^{k-1} \langle \tilde{\varphi}, f \rangle^{k+1} k \frac{(1-u)^{k-1}}{2\varphi(x)} \mathbb{V}[\varphi](x), \quad (52)$$

for which one uses that

$$(k+1)!(\langle \tilde{\varphi}, \beta\mathbb{V}[\varphi] \rangle / 2)^{k-1} \langle \tilde{\varphi}, f \rangle^{k+1} k \mathbb{V}[\varphi](x)$$
$$= \mathcal{E}_x\left[\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \sum_{k_i=1}^{k} \binom{k+1}{k_i, \, k+1-k_i} \varphi(x_i)\varphi(x_j) \right.$$
$$\left. \times \frac{k_i! \, \langle f, \tilde{\varphi} \rangle^{k_i} \, \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k_i-1}}{2^{(k_i-1)}} \frac{(k+1-k_i)! \, \langle f, \tilde{\varphi} \rangle^{k+1-k_i} \, \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-k_i}}{2^{(k-k_i)}} \right].$$

Note that, thanks to the assumption (H2), the expression for $F(s, x)$ clearly satisfies (A.1).

Subtracting the right-hand side of (52) from the right-hand side of (51), again appealing to the induction hypotheses, specifically the second statement in (18), it is

not difficult to show that, for each $\varepsilon \in (0, 1)$,

$$\lim_{t \to \infty} \sup_{x \in E, u \in [0, \varepsilon), f \in B_1^+(E)} |\varphi(x) F[f](x, u, t) - \varphi(x) F[f](x, u)| = 0.$$

On the other hand, the first statement in the induction hypothesis (18) also implies that three exists a constant $C_k > 0$ (which depends on k but not ε) such that

$$\lim_{t \to \infty} \sup_{x \in E, u \in [\varepsilon, 1], f \in B_1^+(E)} |\varphi(x) F(x, u, t) - \varphi(x) F(x, u)| \le C_k (1 - \varepsilon)^{k-1}.$$

Since we may take ε arbitrarily close to 1, we conclude that (A.2) holds.

In conclusion, since the conditions of Theorem A.1 are now met, we get the two statements of (18) as a consequence. $\qquad\square$

2.4 Proof of Theorem 4

Next we turn our attention to the evolution equation generated by the k-th moment functional $\mathrm{M}_t^{(k)}$, $t \ge 0$. To this end, we start by defining observing that

$$\mathrm{M}_t^{(k)}[g](x) = (-1)^{k+1} \frac{\partial^k}{\partial \theta^k} \mathrm{u}_t[0, \theta g](x) \Big|_{\theta = 0}. \tag{53}$$

Taking account of (35), we see that

$$\mathrm{u}_t[0, \theta g](x) = -\int_0^t \mathrm{T}_s \left[\mathrm{A}[\mathrm{u}_{t-s}[0, \theta g]] - \theta g (1 - \mathrm{u}_{t-s}[0, \theta g]) \right](x) \mathrm{d}s. \tag{54}$$

Given the proximity of (54) to (42), it is easy to see that we can apply the same reasoning that we used for $\mathrm{T}_t^{(k)}[f](x)$ to $\mathrm{M}_t^{(k)}[g](x)$ and conclude that, for $k \ge 2$,

$$\mathrm{M}_t^{(k)}[g](x) = \int_0^t \mathrm{T}_s \left[\beta \hat{\eta}_{t-s}^{(k-1)}[g] \right](x) - k \mathrm{T}_s[g \mathrm{M}_{t-s}^{(k-1)}[g]](x) \mathrm{d}s, \tag{55}$$

where $\hat{\eta}^k$ plays the role of η^k albeit replacing the moment operators $\mathrm{T}^{(j)}$ by the moment operators $\mathrm{M}^{(j)}$.

We now proceed to prove Theorem 4, also by induction. First we consider the setting $k = 1$. In that case,

$$\frac{1}{t} \mathrm{M}^{(1)}[g](x) = \frac{1}{t} \mathbb{E}_{\delta_x} \left[\int_0^t \langle g, X_s \rangle \mathrm{d}s \right] = \frac{1}{t} \int_0^t \mathrm{T}_s[g](x) \mathrm{d}s = \int_0^1 \mathrm{T}_{ut}[g](x) \mathrm{d}u.$$

Referring now to Theorem A.1 in the Appendix, we can take $F(x, s, t) = f(x)/\varphi(x)$, since $f \in B^+(E)$, the conditions of the theorem are trivially met and hence

$$\lim_{t \to \infty} \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{t} M^{(1)}[g](x) - \langle g, \tilde{\varphi} \rangle \right| = 0.$$

Note that this limit sets the scene for the polynomial growth in $t^{n(k)}$ of the higher moments for some function $n(k)$. If we are to argue by induction, whatever the choice of $n(k)$, it must satisfy $n(1) = 1$.

Next suppose that Theorem 4 holds for all integer moments up to and including $k-1$. To make the inductive step, we follow similar arguments to the proof of Theorem 1. That is, we use Theorem A.1 and the induction hypothesis to show that the limit of the right-hand side of (55) is precisely the limit that appears in the statement of the theorem.

We have from (55) that

$$\frac{1}{t^{2k-1}} M_t^{(k)}[g](x) = \frac{1}{t^{2k-1}} \int_0^t T_s \left[\beta \hat{\eta}_{t-s}^{(k-1)}[g] \right](x) ds$$
$$- \frac{1}{t^{2k-1}} \int_0^t k T_s [g M_{t-s}^{(k-1)}[g]](x) ds. \qquad (56)$$

Let us first deal with the right most integral in (56). It can be written as

$$\frac{1}{t^{2k-2}} \int_0^1 k T_{ut} [\varphi F(\cdot, u, t)](x) du$$
$$= \int_0^1 (1-u)^{2k-2} k T_{ut} \left[g \frac{1}{(t(1-u))^{2k-2}} M_{t(1-u)}^{(k-1)}[g] \right](x) du.$$

Arguing as in the spirit of the proof of Theorem 1, our induction hypothesis ensures that

$$\lim_{t \to \infty} F[g](x, u, t) = \lim_{t \to \infty} g(1-u)^{2k-2} k \frac{1}{(t(1-u))^{2k-2}} \frac{M_{t(1-u)}^{(k-1)}[g](x)}{\varphi(x)} = 0$$
$$=: F(x, u)$$

satisfies (A.1) and (A.2). Theorem A.1 thus tells us that, uniformly in $x \in E$ and $g \in B_1^+(E)$,

$$\lim_{t \to \infty} \frac{1}{t^{2k-1}} \int_0^t k T_s [g M_{t-s}^{(k-1)}[g]](x) = 0. \qquad (57)$$

Now turning our attention to the first integral on the right-hand side of (56) and again following the style of the reasoning in the proof of Theorem 1, we can pull out

the leading order terms, uniformly for $x \in E$ and $g \in B_1^+(E)$,

$$
\lim_{t \to \infty} \frac{1}{t^{2k-1}} \int_0^t \mathrm{T}_s \left[\beta \hat{\eta}_{t-s}^{(k-1)}[g] \right](x) \mathrm{d}s
$$
$$
= \lim_{t \to \infty} \int_0^1 \mathrm{T}_{ut} \left[\frac{\beta(\cdot)}{2} (1-u)^{2k-2} \mathcal{E}. \left[\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k_i=1}^{k-1} \binom{k}{k_i, \ k - k_i} \varphi(x_i) \varphi(x_j) \right. \right.
$$
$$
\left. \left. \times \frac{\mathrm{M}_{t(1-u)}^{(k_i)}[g](x_i)}{\varphi(x_i)(t(1-u))^{2k_i-1}} \frac{\mathrm{M}_{t(1-u)}^{(k-k_i)}[g](x_j)}{\varphi(x_j)(t(1-u))^{2k-2k_i-1}} \right] \right](x) \mathrm{d}u. \tag{58}
$$

It is again worth noting here that the choice of the polynomial growth in the form $t^{n(k)}$ also constrains the possible linear choices of $n(k)$ to $n(k) = 2k - 1$ if we are to respect $n(1) = 1$ and the correct distribution of the index across (58).

Identifying

$$
F[g](x, u, t) = \frac{\beta(x)}{2\varphi(x)} (1-u)^{2k-2} \mathcal{E}_x \left[\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k_i=1}^{k-1} \binom{k}{k_i, \ k - k_i} \varphi(x_i) \varphi(x_j) \right.
$$
$$
\left. \times \frac{\mathrm{M}_{t(1-u)}^{(k_i)}[g](x_i)}{\varphi(x_i)(t(1-u))^{2k_i-1}} \frac{\mathrm{M}_{t(1-u)}^{(k-k_i)}[g](x_j)}{\varphi(x_j)(t(1-u))^{2k-2k_i-1}} \right],
$$

our induction hypothesis allows us to conclude that $F[g](x, u) := \lim_{t \to \infty} F[g](x, u, t)$ exists and

$$
\varphi(x) F[g](x, u) = (1-u)^{2k-2} k! \frac{\beta(x) \mathbb{V}[\varphi](x)}{2^{k-1}} \langle g, \tilde{\varphi} \rangle^k \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \sum_{\ell=1}^{k-1} L_\ell L_{k-\ell}.
$$

Thanks to our induction hypothesis, we can also easily verify (A.1) and (A.2). Theorem A.1 now gives us the required uniform (in $x \in E$ and $g \in B_1^+(E)$) limit

$$
\lim_{t \to \infty} \frac{1}{t^{2k-1}} \int_0^t \mathrm{T}_s \left[\beta \hat{\eta}_{t-s}^{(k-1)}[g] \right](x) \mathrm{d}s = \frac{k! \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \langle g, \tilde{\varphi} \rangle^k}{2^{k-1}} L_k. \tag{59}
$$

Putting (59) together with (57) we get the statement of Theorem 4. $\qquad \square$

2.5 Remaining proofs in the non-critical cases

We now give an outline of the main steps in the proof of Theorem 1 for the sub and supercritical cases. As previously mentioned, the ideas used in this section will closely follow those presented in the previous section for the proof of the critical case and so we leave the details to the reader. We first note that the Perron Frobenius behaviour

in (H1) ensures the base case for the induction argument, regardless of the value of λ. We thus turn to the inductive step, assuming the result holds for $k - 1$.

Proof of Theorem 2 (supercritical case) From the evolution equation (41), we have

$$
\lim_{t \to \infty} \frac{e^{-\lambda k t}}{\varphi(x)} T_t^{(k)}[f](x)
$$

$$
= \lim_{t \to \infty} \frac{e^{-\lambda k t}}{\varphi(x)} \int_0^t T_s \left[\beta \mathcal{E}_x \left[\sum_{[k_1, \ldots, k_N]_k^2} \binom{k}{k_1, \ldots, k_N} \prod_{j=1}^N T_{t-s}^{(k_j)}[f](x_j) \right] \right](x) \mathrm{d}s.
$$
(60)

It then follows that

$$
\lim_{t \to \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} e^{-k\lambda t} T_t^{(k)}[f](x) - \langle f, \tilde{\varphi} \rangle^k L_k \right|
$$

$$
= \lim_{t \to \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} e^{-k\lambda t} \int_0^t T_s \right.
$$

$$
\left[\beta \mathcal{E}_x \left[\sum_{[k_1, \ldots, k_N]_k^2} \binom{k}{k_1, \ldots, k_N} \prod_{j=1}^N T_{t-s}^{(k_j)}[f](x_j) \right] \right](x) \mathrm{d}s
$$

$$
\left. - \int_0^t e^{-(k-1)\lambda s} \mathrm{d}s \, \langle f, \tilde{\varphi} \rangle^k \lambda (k-1) L_k \right|.
$$
(61)

Noting that $\sum_{j=1}^N k_j = k$, we may again share the exponential term across the product in the right-hand side above as follows,

$$
\frac{e^{-\lambda k t}}{\varphi(x)} \int_0^t T_s \left[\beta \mathcal{E} \left[\sum_{[k_1, \ldots, k_N]_k^2} \binom{k}{k_1, \ldots, k_N} \prod_{j=1}^N T_{t-s}^{(k_j)}[f](x_j) \right] \right](x) \mathrm{d}s
$$

$$
= t \int_0^1 e^{-\lambda(k-1)ut} \frac{e^{-\lambda ut}}{\varphi(x)} T_{ut}
$$

$$
\left[k! \beta \mathcal{E} \left(\sum_{[k_1, \ldots, k_N]_k^2} \prod_{j=1}^N \varphi(x_j) \frac{e^{-\lambda k_j t(1-u)} T_{t(1-u)}^{(k_j)}[f](x_j)}{k_j! \varphi(x_j)} \right) \right](x) \mathrm{d}u.
$$

Combining this with to (61) and changing variables in the final integral of the latter, we have

$$\lim_{t\to\infty} \sup_{x\in E, f\in B_1^+(E)} \left| \varphi(x)^{-1} e^{-k\lambda t} T_t^{(k)}[f](x) - \langle f, \tilde{\varphi} \rangle^k L_k \right|$$

$$\leq \lim_{t\to\infty} \sup_{x\in E, f\in B_1^+(E)} t \left| \int_0^1 e^{-\lambda(k-1)ut} \left(\frac{e^{-\lambda ut}}{\varphi(x)} T_{ut}[\varphi F(\cdot, u, t)] - \langle f, \tilde{\varphi} \rangle^k \lambda(k-1) L_k \right) du \right|,$$
(62)

where we have defined

$$F[f](x, u, t) := k! \frac{\beta(x)}{\varphi(x)} \mathcal{E}_x \left[\sum_{[k_1,\dots,k_N]_k^2} \prod_{j=1}^N \varphi(x_j) \frac{e^{-\lambda k_j t(1-u)} T_{t(1-u)}^{(k_j)}[f](x_j)}{k_j! \varphi(x_j)} \right].$$

It is easy to see that, pointwise in $x \in E$ and $u \in [0, 1]$, using the induction hypothesis and (H2),

$$F[f](x, u) := \lim_{t\to\infty} F[f](x, u, t)$$

$$= k! \frac{\beta(x)}{\varphi(x)} \mathcal{E}_x \left[\sum_{[k_1,\dots,k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{\substack{j=1 \\ j:k_j>0}}^N \varphi(x_j) L_{k_j} \right] \langle f, \tilde{\varphi} \rangle^k,$$

where we have again used the fact that the k_js sum to k to extract the $\langle f, \tilde{\varphi} \rangle^k$ term. Similarly to the critical setting we can also verify using the induction hypothesis and (H2) that (A.1) and (A.2) hold.

This is sufficient to note that, by using a triangle inequality similar spirit to the one found in (A.4) and appealing to (14) of the assumption (H1), we have that

$$\sup_{x\in E, u\in[0,1], f\in B_1^+(E), t\geq 0} \left| \frac{e^{-\lambda ut}}{\varphi(x)} T_{ut}[\varphi F[f](\cdot, u, t)] - k! \langle f, \tilde{\varphi} \rangle^k L_k \right| < \infty.$$

This means that for t sufficiently large, we can control the modulus in the integral on the right-hand side of (62) by a global constant. The remainder of integral, yields a bound of $\varepsilon(1 - e^{-\lambda(k-1)t})/\lambda(k - 1)$, which tends to zero as $t \to \infty$. □

Proof of Theorem 3 (subcritical case) We now outline the subcritical case. First note that since we only compensate by $e^{-\lambda t}$, the term $T_t[f^k](x)$ that appears in equation (41) does not vanish after the normalisation. Due to assumption (H1), we have

$$\lim_{t\to\infty} \varphi^{-1}(x) e^{-\lambda t} T_t[f^k](x) = \langle f^k, \tilde{\varphi} \rangle.$$

Next we turn to the integral term in (41). Define $[k_1, \dots, k_N]_k^n$, for $2 \leq n \leq k$ to be the set of tuples $(k_1, \dots, k_N)$ with exactly n positive terms and whose sum is equal to

k. Similar calculations to those given above yield

$$\frac{e^{-\lambda t}}{\varphi(x)} \int_0^t T_s \left[\beta \mathcal{E}_x \left[\sum_{[k_1,\ldots,k_N]_k^2} \binom{k}{k_1,\ldots,k_N} \prod_{j=1}^N T_{t-s}^{(k_j)}[f](x_j) \right] \right](x) ds$$

$$= t \sum_{n=2}^k \int_0^1 e^{\lambda(n-1)ut} \frac{e^{-\lambda t(1-u)}}{\varphi(x)} T_{t(1-u)} \left[k! \beta \mathcal{E}. \right.$$

$$\left. \left[\sum_{[k_1,\ldots,k_N]_k^n} \prod_{j=1}^N \varphi(x_j) \frac{e^{-\lambda ut} T_{ut}^{(k_j)}[f](x_j)}{k_j! \varphi(x_j)} \right] \right](x) du \tag{63}$$

Again, we leave the details to the reader but the idea is that the induction hypothesis will take care of the product of the lower order moments and the second part of (14) in assumption (H1) will then take care of the asymptotic behaviour semigroup $T_{t(1-u)}$. The second part of (14) allows one to control the difference between this term and its limit. In a similar manner to the final step in the proof of Theorem 2, the difference of (41) and its limit can be reduced to the limit as $t \to \infty$ of $\varepsilon(1 - e^{-|\lambda|(n-1)t})/|\lambda|(n-1)$, which is bounded above by ε. □

Proof of Theorem 5 For the case $k = 1$, we have

$$\left| e^{-\lambda t} \varphi(x)^{-1} \int_0^t T_s[g](x) ds - \frac{\langle g, \tilde\varphi \rangle}{\lambda} \right|$$

$$= \left| e^{-\lambda t} t \int_0^1 e^{\lambda ut} \left(e^{-\lambda ut} \varphi(x)^{-1} T_{ut}[g](x) - \langle g, \tilde\varphi \rangle \right) du - e^{-\lambda t} \frac{\langle g, \tilde\varphi \rangle}{\lambda} \right|$$

$$\leq e^{-\lambda t} t \int_0^1 e^{\lambda ut} \left| e^{-\lambda ut} \varphi(x)^{-1} T_{ut}[g](x) - \langle g, \tilde\varphi \rangle \right| du + e^{-\lambda t} \frac{\langle g, \tilde\varphi \rangle}{\lambda}. \tag{64}$$

Thanks to (H1) and similar arguments to those used in the proof of Theorem 2, we may choose t sufficiently large such that the modulus in the integral is bounded above by $\varepsilon > 0$, uniformly in $g \in B_1^+(E)$ and $x \in E$. Then, the right-hand side of (64) is bounded above by $\varepsilon \lambda^{-1}(1 - e^{-\lambda t}) + e^{-\lambda t} \langle g, \tilde\varphi \rangle / \lambda$. Since ε can be taken arbitrarily small, this gives the desired result and also pins down the initial value $L_1 = 1/\lambda$.

Now assume the result holds for all $\ell \leq k-1$. Reflecting on proof of Theorem 2, we note that in this setting the starting point is almost identical except that the analogue of (60), which is derived from (55), is now the need to evaluate

$$\lim_{t\to\infty} \frac{e^{-\lambda k t}}{\varphi(x)} M_t^{(k)}[g](x) = \lim_{t\to\infty} \frac{e^{-\lambda k t}}{\varphi(x)} \int_0^t T_s$$

$$\left[\beta \mathcal{E}. \left[\sum_{[k_1,\ldots,k_N]_k^2} \binom{k}{k_1,\ldots,k_N} \prod_{j=1}^N M_{t-s}^{(k_j)}[g](x_j) \right] \right](x) ds$$

$$- k \lim_{t \to \infty} \frac{e^{-\lambda k t}}{\varphi(x)} \int_0^t T_s[g M_{t-s}^{(k-1)}[g]](x) ds. \tag{65}$$

The first term on the right-hand side of (65) can be handled in essentially the same way as in the proof of Theorem 2. The second term on the right-hand side of (65) can easily be dealt with along the lines that we are now familiar with from earlier proofs, using the induction hypothesis. In particular, its limit is zero. Hence combined with the first term on the right-hand side of (65), we recover the same recursion equation for L_k. $\qquad\square$

Proof of Theorem 6 The case $k = 1$ is relatively straightforward and, again in the interest of keeping things brief, we point the reader to the fact that, as $t \to \infty$, we have

$$\frac{1}{\varphi(x)} M_t^{(1)}[g](x) = \int_0^t \frac{T_s[g](x)}{\varphi(x)} ds$$
$$= \int_0^t e^{\lambda s} e^{-\lambda s} \frac{T_s[g](x)}{\varphi(x)} ds \sim t \langle g, \tilde{\varphi} \rangle \int_0^1 e^{\lambda u t} du \sim \frac{\langle g, \tilde{\varphi} \rangle}{|\lambda|}. \tag{66}$$

Now suppose the result holds for all $\ell \leq k - 1$. We again refer to (55), which means we are interested in handling a limit which is very similar to (65), now taking the form

$$\frac{M_t^{(k)}[g](x)}{\varphi(x)}$$
$$= \frac{t}{\varphi(x)} \int_0^1 e^{\lambda u t} e^{-\lambda u t} T_{ut}$$
$$\left[\beta \mathcal{E}. \left[\sum_{[k_1, \ldots, k_N]_k^2} \binom{k}{k_1, \ldots, k_N} \prod_{j=1}^N \varphi(x_j) \frac{M_{t(1-u)}^{(k_j)}[g](x_j)}{\varphi(x_j)} \right] \right](x) du$$
$$- k \frac{t}{\varphi(x)} \int_0^1 e^{\lambda u t} e^{-\lambda u t} T_{ut} \left[g \varphi \frac{M_{t(1-u)}^{(k-1)}[g]}{\varphi} \right](x) du. \tag{67}$$

Again skipping the details, we can quickly see from (67) the argument in (66), and the induction hypothesis gives us

$$\frac{M_t^{(k)}[g](x)}{\varphi(x)} \sim k! \frac{\langle g, \tilde{\varphi} \rangle^k}{|\lambda|} \left\langle \beta \mathcal{E} \left[\sum_{\substack{[k_1, \ldots, k_N]_k^2 \\ j : k_j > 0}} \prod_{\substack{j=1 \\ j : k_j > 0}}^N \varphi(x_j) L_{k_j} \right], \tilde{\varphi} \right\rangle$$
$$- k! \frac{\langle g\varphi, \tilde{\varphi} \rangle \langle g, \tilde{\varphi} \rangle^{k-1}}{|\lambda|} L_{k-1}, \tag{68}$$

which gives us the required recursion for L_k. $\qquad\square$

3 Proofs for superprocesses

For the proof of Theorems 1, 2 and 3 in the setting of superprocesses we follow a similar approach. One difference is that we cannot work with the k-th moment as a product of an almost surely finite sum. As such the use of the Leibniz formula as in the previous section is no longer helpful. Instead, we use the Faà di Bruno formula (see Lemma A.1) to assist with multiple derivatives of the non-linear evolution equation (5).

3.1 Linear and non-linear semigroup equations

The evolution equation for the expectation semigroup $(\mathrm{T}_t, t \geq 0)$ is well known and satisfies

$$\mathrm{T}_t\,[f]\,(x) = \mathrm{P}_t[f](x) + \int_0^t \mathrm{P}_s\left[\beta(\mathrm{m}[\mathrm{T}_{t-s}[f]] - 1) + b\right](x)\mathrm{d}s, \qquad (69)$$

for $t \geq 0$, $x \in E$ and $f \in B^+(E)$, where, with a meaningful abuse of our branching Markov process notation, we now define

$$\mathrm{m}[f](x) = \int_{M_0(E)}\left[\gamma(x, \pi)\,\langle f, \pi\rangle + \int_0^\infty u\,\langle f, \pi\rangle\,n(x, \pi, \mathrm{d}u)\right] G(x, \mathrm{d}\pi)$$

$$= \gamma(x, f) + \int_{M(E)^\circ}\langle f, \nu\rangle\Gamma(x, \nu). \qquad (70)$$

See for example equation (3.24) of [7].

In the spirit of Lemma 1 we can give a second representation of $\mathrm{T}_t[f]$ in terms of an auxiliary process, the so called many-to-one formula. To this end, if, as before, we work with the process $(\xi, \mathbf{P})$ to represent the Markov process associated to the semigroup $(\mathrm{P}_t, t \geq 0)$, then, although we have redefined the quantity $\mathrm{m}[f](x)$, we can still meaningfully work with the process $(\hat{\xi}, \hat{\mathbf{P}})$ as defined just before Lemma 1.

Lemma 3 *Let* $\vartheta(x) = B(x) + b(x) = \beta(x)(\mathrm{m}[1](x) - 1) + b(x)$, *then, for* $t \geq 0$ *and* $f \in B^+(E)$,

$$\mathrm{T}_t\,[f]\,(x) = \hat{\mathbf{E}}_x\left[\exp\left(\int_0^t \vartheta(\hat{\xi}_s)\mathrm{d}s\right) f(\hat{\xi}_t)\right]. \qquad (71)$$

As with Lemma 1, the proof is classical, requiring only that we take the right-hand side of (71) and condition on the first extra jump of $(\hat{\xi}, \hat{\mathbf{P}})$ to show that it also solves (70). It is a straightforward application of Grönwall's inequality to show that (70) has a unique solution and hence (69) holds. The reader will note that because we have separated out the local and non-local branching mechanisms of the superprocess, the deliberate repeat definition of $\mathrm{m}[f]$ for superprocesses is only the analogue of its counter part for branching Markov processes in the sense of non-local activity. The mean local branching rate has otherwise been singled out as the term b.

Similarly to the branching Markov process setting, let us re-write an extended version of the non-linear semigroup evolution $(V_t, t \geq 0)$, defined in (6), i.e. the natural analogue of (35), in terms of the linear semigroup $(T_t, t \geq 0)$. To this end, define

$$V_t[f, g](x) = \mathbb{E}_x \left[e^{-\langle f, X_t \rangle - \int_0^t \langle g, X_s \rangle ds} \right],$$

Analogously to Theorem 2 we have the following result.

Lemma 4 *For all $f, g \in B^+(E)$, $x \in E$ and $t \geq 0$, the non-linear semigroup $V_t[f, g](x)$ satisfies*

$$V_t[f, g](x) = T_t[f](x) - \int_0^t T_s \left[\mathcal{J}[V_{t-s}[f, g]] - g V_{t-s}[f, g] \right](x) ds, \quad (72)$$

where, for $h \in B^+(E)$ and $x \in E$,

$$\mathcal{J}[h](x) = \psi(x, h(x)) + \phi(x, h) + \beta(x)(m[h](x) - h(x)) + b(x)h(x).$$

The proof is essentially the same as the proof of Lemma 2 and hence we leave the details to the reader.

3.2 Evolution equations for the *k*-th moment of a superprocesses

Recall that we defined $T_t^{(k)}[f](x) := \mathbb{E}_{\delta_x}[\langle f, X_t \rangle^k]$, $t \geq 0$, $f \in B^+(E)$, $k \geq 1$. As with the setting of branching Markov processes, we want to establish an evolution equation for $(T_t^{(k)}, t \geq 0)$, from which we can establish the desired asymptotics. To this end, let us introduce the following notation.

For $x \in E$, $k \geq 2$ and $t \geq 0$, define

$$R_k(x, t) = \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} (-1)^{m_1 + \dots + m_{k-1} - 1}$$

$$(m_1 + \dots + m_{k-1} - 1)! \prod_{j=1}^{k-1} \left(\frac{(-1)^j T_t^{(j)}[f](x)}{j!} \right)^{m_j}, \quad (73)$$

and

$$K_k(x, t) = \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} \psi^{(m_1 + \dots + m_{k-1})}(x, 0+)$$

$$\prod_{j=1}^{k-1} \left(\frac{(-1)^{j+1} T_t^{(j)}[f](x) - R_j(x, t)}{j!} \right)^{m_j}, \quad (74)$$

and finally

$$
S_k(x, t) = \int_{M(E)^\circ} \sum_{\{m_1,\dots,m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} (-1)^{m_1+\cdots+m_{k-1}}
$$
$$
\prod_{j=1}^{k-1} \left(\frac{\left\langle (-1)^{j+1} T_t^{(j)} [f] - R_j(\cdot, t), v \right\rangle}{j!} \right)^{m_j} \Gamma(x, dv), \qquad (75)
$$

and the sums run over the set of non-negative integers $\{m_1, \dots, m_{k-1}\}$ such that $m_1 + 2m_2 + \cdots + (k-1)m_{k-1} = k$.

Theorem 7 *Fix $k \geq 2$. Suppose that (H1) and (H2) hold, with the additional assumption that*

$$
\sup_{x \in E, s \leq t} T_s^{(\ell)}[f](x) < \infty, \qquad \ell \leq k - 1, \, f \in B^+(E), \, t \geq 0. \qquad (76)
$$

Then,

$$
T_t^{(k)}[f](x) = (-1)^{k+1} R_k(x, t) + (-1)^k \int_0^t T_s[U_k(\cdot, t - s)] \, ds, \qquad (77)
$$

where

$$
U_k(x, t) = K_k(x, t) + \beta(x) S_k(x, t). \qquad (78)
$$

Proof First note that similarly to the Markov branching process case, defining

$$
e_t[f](x) := \mathbb{E}_{\delta_x} \left[e^{-\langle f, X_t \rangle} \right], \qquad t \geq 0, \, f \in B^+(E),
$$

we have

$$
e_t^{(k)}[\theta f](x) := \frac{\partial^k}{\partial \theta^k} e_t[\theta f](x) = (-1)^k \mathbb{E}_{\delta_x} \left[\langle f, X_t \rangle^k e^{-\theta \langle f, X_t \rangle} \right]
$$

and

$$
e_t^{(k)}[\theta f](x)|_{\theta=0} = (-1)^k T_t^{(k)}[f](x). \qquad (79)
$$

To prove (77), recall the definition (5) and let

$$
v_t^{(k)}[f](x) := \frac{\partial^k}{\partial \theta^k} V_t[\theta f, 0](x) \Big|_{\theta=0}
$$
$$
= -\frac{\partial^k}{\partial \theta^k} \log e_t[\theta f](x) \Big|_{\theta=0}, \qquad t \geq 0, \, f \in B^+(E), \, k \geq 1
$$

The idea is to use Lemma A.1 to obtain two equivalent expressions for $v_t^{(k)}[f](x)$ that, when equated, yield (77).

To this end, not that due to Faà di Bruno's Lemma A.1, we have

$$
v_t^{(k)}[f](x)
$$

$$
= \frac{\partial^k}{\partial\theta^k} - \log e_t[\theta f](x)\Big|_{\theta=0}
$$

$$
= - \sum_{\{m_1,\dots,m_k\}_k} \frac{k!}{m_1!\dots m_k!} \frac{(-1)^{m_1+\cdots+m_k-1}(m_1+\cdots+m_k-1)!}{e_t[\theta f](x)^{m_1+\cdots+m_k}}
$$

$$
\prod_{j=1}^{k}\left(\frac{e^{(j)}[\theta f])}{j!}\right)^{m_j}\Bigg|_{\theta=0}
$$

$$
= - \sum_{\{m_1,\dots,m_k\}_k} \frac{k!}{m_1!\dots m_k!}(-1)^{m_1+\cdots+m_k-1}(m_1+\cdots+m_k-1)!
$$

$$
\prod_{j=1}^{k}\left(\frac{(-1)^j T_t^{(j)}[f](x)}{j!}\right)^{m_j},
$$

where the sum runs over the set of non-negative integers $\{m_1,\dots,m_k\}_k$ such that

$$
m_1 + 2m_2 + \cdots + km_k = k.
$$

Note that $m_k > 0$ if and only if $m_k = 1$ and $m_1 = m_2 = \cdots = m_{k-1} = 0$, so the k-th moment term $T_t^{(k)}[f]$ appears only once and with a factor $(-1)^{k+1}$, that is,

$$
v_t^{(k)}[f](x) = (-1)^{k+1} T_t^{(k)}[f](x) - R_k(x,t), \tag{80}
$$

where all the terms in $R_k(x,t)$ are products of two or more lower order moments. Thus it remains to show that

$$
v_t^{(k)}[f](x) = - \int_0^t T_s\left[U_k(\cdot, t-s)\right]ds.
$$

Differentiating the evolution equation (72) k times at $\theta = 0$, momentarily not worrying about passing derivatives through integrals, we get

$$
v_t^{(k)}[f](x) = - \int_0^t T_s\left[\frac{\partial^k}{\partial\theta^k}\Big(\psi(\cdot, V_{t-s}[\theta f, 0](\cdot))\right.
$$

$$
\left. + \phi(\cdot, V_{t-s}[\theta f, 0]) + F[V_{t-s}[\theta f, 0]]\Big)\Big|_{\theta=0}\right](x)ds,
$$

where

$$\mathrm{F}[g](x) = \beta(x)(m[g] - g) + b(x)g, \qquad x \in E, g \in B^+(E).$$

We first deal with the k-th derivative of the term involving ψ in the above integral. For this, we again use Lemma A.1 to get

$$\frac{\partial^k}{\partial \theta^k} \psi\left(x, \mathrm{V}_t\left[\theta f, 0\right](x)\right)\bigg|_{\theta=0}$$

$$= \sum_{\{m_1,\ldots,m_k\}_k} \frac{k!}{m_1! \ldots m_k!} \psi^{(m_1+\cdots+m_k)}(x, \mathrm{V}_t\left[\theta f, 0\right])$$

$$\prod_{j=1}^{k} \left(\frac{\frac{\partial^j}{\partial \theta^j} \mathrm{V}_t\left[\theta f, 0\right](x)}{j!}\right)^{m_j}\bigg|_{\theta=0}$$

$$= \sum_{\{m_1,\ldots,m_k\}_k} \frac{k!}{m_1! \ldots m_k!} \psi^{(m_1+\cdots+m_k)}(x, 0+) \prod_{j=1}^{k} \left(\frac{\mathrm{v}_t^{(j)}[f](x)}{j!}\right)^{m_j}$$

$$= -b(x)\mathrm{v}_t^{(k)}[f](x) + K_k(x, t),$$

where the last equality holds because $m_k = 1$ if and only if $m_1 = \cdots = m_{k-1} = 0$ and $\psi'(x, 0+) = -b(x)$.

Similarly, for the the kth derivative of the remaining terms, recalling (8), (9) and (70), we have

$$\frac{\partial^k}{\partial \theta^k} \left(\phi(x, \mathrm{V}_t\left[\theta f, 0\right]) + \mathrm{F}[\mathrm{V}_t\left[\theta f, 0\right]]\right)$$

$$= b(x)\frac{\partial^k}{\partial \theta^k} \mathrm{V}_t\left[\theta f, 0\right]$$

$$- \beta(x) \int_{M_0(E)} \int_0^\infty \frac{\partial^k}{\partial \theta^k} \left(1 - e^{-u\langle \mathrm{V}_t[\theta f, 0], \pi\rangle} - u\langle \mathrm{V}_t\left[\theta f, 0\right], \pi\rangle\right)$$

$$n(x, \pi, \mathrm{d}u)G(x, \mathrm{d}\pi).$$

Using Lemma A.1 yields

$$\frac{\partial^k}{\partial \theta^k} \left(1 - e^{-u\langle \mathrm{V}_t[\theta f, 0], \pi\rangle} - u\langle \mathrm{V}_t\left[\theta f, 0\right], \pi\rangle\right)$$

$$= \sum_{\{m_1,\ldots,m_{k-1}\}_k} \frac{k!}{m_1! \ldots m_k!} (-1)^{m_1+\cdots+m_k+1} e^{-u\langle \mathrm{V}_t[\theta f, 0], \pi\rangle}$$

$$\prod_{j=1}^{k} \left(\frac{u \left\langle \frac{\partial^j}{\partial \theta^j} V_t \left[\theta f, 0 \right], \pi \right\rangle}{j!} \right)^{m_j}$$

$$+ \left(e^{-u \langle V_t [\theta f, 0], \pi \rangle} - 1 \right) u \left\langle \frac{\partial^k}{\partial \theta^k} V_t \left[\theta f, 0 \right], \pi \right\rangle,$$

where, in the final equality, we have singled out the case that $m_k = 1$ and $m_1 = \cdots = m_{k-1} = 0$ in the Faà di Bruno formula. Using the definition of $m[f](x)$ in (70) and the same observation as above about the m_j's, we get

$$\frac{\partial^k}{\partial \theta^k} \left(\phi(x, V_t [\theta f, 0]) + F[V_t [\theta f, 0]] \right) \bigg|_{\theta = 0} = b(x) v_t^{(k)}[f](x) + \beta(x) S_k(x, t). \quad (81)$$

Putting the pieces together, we obtain

$$v_t^{(k)}[f](x) = - \int_0^t T_s \left[U_k(\cdot, t - s) \right](x) \mathrm{d}s. \quad (82)$$

Combining this with equation (80) yields

$$(-1)^{k+1} T_t^{(k)} [f](x) = R_k(x, t) - \int_0^t T_s \left[U_k(\cdot, t - s) \right](x) \mathrm{d}s,$$

which is the desired result.

There is one final matter we must attend to, which is the ability to move derivatives through integrals. In this setting, this follows from the assumption (76), (H2) and the Lévy-Khintchine-type formulae for ψ and ϕ. □

3.3 Completing the proof of Theorem 1: critical case

We will prove Theorem 1 for superprocesses using induction, similarly to the setting of branching Markov processes. The case $k = 1$ follows from assumption (H1).

Now assume that the statement of Theorem 1 holds in the superprocess setting for all $\ell \leq k$. Our aim is to prove that the result holds for $k + 1$. Using Theorem 7 and a change of variables, we have that

$$\frac{1}{\varphi(x) t^k} T_t^{(k+1)} [f](x) = \frac{(-1)^k}{\varphi(x) t^k} R_{k+1}(x, t)$$

$$+ \frac{(-1)^{k+1}}{\varphi(x) t^{k-1}} \int_0^1 T_{st} \left[U_{k+1}(\cdot, t(1 - s)) \right](x) \mathrm{d}s, \quad (83)$$

where R and U were defined in equations (73) and (78), respectively. As with the particle system, we first aim to simplify the right-hand side before showing that its limit is equivalent to the expression given in the statement of the theorem. In particular, we

will first show that for each $x \in E$, the limit of the right-hand side of (83) is equivalent to

$$\lim_{t \to \infty} \frac{1}{2\varphi(x)t^{k-1}} \int_0^1 \mathrm{T}_{st} \left[K_{k+1}^{(2)}(\cdot, t(1-s)) + \beta(\cdot) S_{k+1}^{(2)}(\cdot, t(1-s)) \right] (x) \mathrm{d}s, \quad (84)$$

where

$$K_{k+1}^{(2)}(x, t) := \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1! k_2!} \psi''(x, 0+) \mathrm{T}_t^{(k_1)}[f](x) \mathrm{T}_t^{(k_2)}[f](x) \quad (85)$$

and

$$S_{k+1}^{(2)}(x, t) = \int_{M(E)^\circ} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1! k_2!} \langle \mathrm{T}_t^{(k_1)}[f], v \rangle \langle \mathrm{T}_t^{(k_2)}[f], v \rangle \Gamma(x, \mathrm{d}v), \quad (86)$$

such that $\{k_1, k_2\}^+$ is defined to be the set of positive integers k_1, k_2 such that $k_1 + k_2 = k + 1$.

To this end, writing $c(m_1, \ldots, m_k)$ for the constants preceding the product summands in (73), observe that

$$\lim_{t \to \infty} \frac{1}{t^k} R_{k+1}(x, t)$$

$$= \lim_{t \to \infty} \frac{(k+1)!}{t^k} \sum_{\{m_1, \ldots, m_k\}_{k+1}} c(m_1, \ldots, m_k) \prod_{j=1}^k \left(\frac{(-1)^j \mathrm{T}_t^{(j)}[f](x)}{j!} \right)^{m_j}$$

$$= (-1)^k (k+1)! \lim_{t \to \infty} \sum_{\{m_1, \ldots, m_k\}_{k+1}} \frac{c(m_1, \ldots, m_k)}{t^{m_1 + \cdots + m_k - 1}} \prod_{j=1}^k \left(\frac{1}{j!} \frac{\mathrm{T}_t^{(j)}[f](x)}{t^{j-1}} \right)^{m_j} = 0,$$

where the final equality is due to the induction hypothesis and the fact that $m_1 + \cdots + m_k > 1$, which follows from the fact that $m_1 + 2m_2 + \cdots + \cdots + k m_k = k + 1$. Note, moreover that the induction hypothesis ensures that the limit is uniform in $x \in E$ and, in fact, that

$$\sup_{t \geq 0, x \in E} \frac{1}{t^{\ell-1}} R_\ell(x, t) < \infty \quad \text{and} \quad \lim_{t \to \infty} \sup_{x \in E} \frac{1}{t^{\ell-1}} R_\ell(x, t) = 0 \quad \ell = 1, \ldots, k+1. \tag{87}$$

We now return to (83), to deal with the term involving U_{k+1}, which we recall is a linear combination of K_{k+1} and S_{k+1}, which were defined in (73) and (75), respectively. Note that if any of the summands in either K_{k+1} or S_{k+1} have more than two of the m_j positive, the limit of that summand, when renormalised by $1/t^{k-1}$, will be zero. In essence, the argument here is analogous to those that led to (49) in the branching Markov process setting. This implies that the only terms in the sums of

(73) and (75) that remain in the limit of (83) are those for which $m_{k_1} = m_{k_2} = 1$ and $m_j = 0$ for all $j \neq k_1, k_2$, with $k_1 < k_2$ such that $k_1 + k_2 = k + 1$, and if $k + 1$ is even, the terms in which $m_{(k+1)/2} = 2$ and $m_j = 0$ for all $j \neq (k+1)/2$.

Let us now convert all of the above heuristics into rigorous computation, for which we appeal to Theorem A.1. We write

$$F[f](x, s, t) := \frac{1}{\varphi(x) t^{k-1}} \left(K_{k+1}^{(3+)}(x, t(1-s)) + \beta(x) S_{k+1}^{(3+)}(x, t(1-s)) \right), \quad (88)$$

where $K_{k+1}^{(3+)}$ and $S_{k+1}^{(3+)}$ contain the terms in K_{k+1} and S_{k+1}, respectively, for which the sum $m_1 + \cdots + m_k$ is greater than or equal to 3. We will prove that $\lim_{t \to \infty} F(x, s, t) = 0$ and that (A.1) and (A.2) hold.

Due to (16) and boundness of φ, dominated convergence implies that

$$\lim_{t \to \infty} \frac{1}{\varphi(x) t^{k-1}} K_{k+1}^{(3+)}(x, t(1-s))$$

$$= \frac{(k+1)!}{\varphi(x)} \sum_{\{m_1, \ldots, m_k\}_{k+1}^3} \frac{\psi^{(m_1 + \cdots + m_k)}(x, 0+)}{m_1! \ldots m_k!}$$

$$\lim_{t \to \infty} \frac{1}{t^{m_1 + \cdots + m_k - 2}} \prod_{j=1}^{k} \left(\frac{(-1)^{j+1} T_{t(1-s)}^{(j)} [f](x) - R_j(x, t(1-s))}{j! t^{j-1}} \right)^{m_j},$$

where the set $\{m_1, \ldots, m_k\}_{k+1}^3$ is the subset of $\{m_1, \ldots, m_k\}_{k+1}$ for which $m_1 + \cdots + m_k \geq 3$. Using the induction hypothesis and (87), we get that the right-hand side above is zero. The same arguments also imply that the limit of $S_{k+1}^{(3+)}$ is zero. Thus $F[f](x, s) := \lim_{t \to \infty} F[f](x, s, t) = 0$. The condition (A.1) trivially holds. For (A.2), the required uniformity follows from the induction hypothesis and (87).

Using Theorem A.1 in the Appendix, we conclude that

$$\lim_{t \to \infty} \sup_{x \in E} \left| \frac{1}{\varphi(x) t^{k-1}} \int_0^1 T_{st} \left[\varphi F[f](\cdot, s, t) \right](x) \mathrm{d}s \right| = 0.$$

Let us now define $\{k_1 < k_2\}$ to be the elements in K_{k+1} for which $m_{k_1} = m_{k_2} = 1$ with $k_1 < k_2$ such that $k_1 + k_2 = k + 1$ and $m_j = 0$ for all other indices and, in the case where $k + 1$ is even, $m_{(k+1)/2} = 2$ and $m_j = 0$ for all $j \neq (k+1)/2$. Restricting the sum to this set in K_{k+1} we get the following expression

$$K_{k+1}^{(2)}(x, t) = \sum_{\{k_1 < k_2\}} \frac{(k+1)!}{k_1! k_2!} \psi''(x, 0+)(-1)^{k+1} T_t^{(k_1)} [f](x) T_t^{(k_2)} [f](x)$$

$$+ \mathbf{1}_{(k+1 \text{ is even})} \frac{1}{2} \binom{k+1}{k/2} \psi''(x, 0+)(-1)^{k+1} \left(T_t^{(k/2)} [f](x) \right)^2$$

$$= \frac{(-1)^{k+1}}{2} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1! k_2!} \psi''(x, 0+) T_t^{(k_1)} [f](x) T_t^{(k_2)} [f](x),$$

where we recall $\{k_1, k_2\}^+$ is the set of positive integers k_1, k_2 such that $k_1 + k_2 = k+1$. Similarly, we obtain the following expression for S_{k+1}:

$$S_{k+1}^{(2)}(x,t) = \int_{M(E)^\circ} \sum_{k_1 < k_2} \frac{(k+1)!}{k_1! k_2!} (-1)^{k+1} \langle \mathrm{T}_t^{(k_1)}[f], \nu \rangle \langle \mathrm{T}_t^{(k_2)}[f], \nu \rangle \Gamma(x, d\nu)$$

$$+ \mathbf{1}_{(k+1 \text{ is even})} \int_{M(E)^\circ} \binom{k+1}{k/2} \frac{(-1)^{k+1}}{2} \langle \mathrm{T}_t^{((k+1)/2)}[f], \nu \rangle^2 \Gamma(x, d\nu)$$

$$= \frac{(-1)^{k+1}}{2} \int_{M(E)^\circ} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1! k_2!} \langle \mathrm{T}_t^{(k_1)}[f], \nu \rangle \langle \mathrm{T}_t^{(k_2)}[f], \nu \rangle \Gamma(x, d\nu).$$

This shows that the right-hand side of (83) is equivalent to (84).

To conclude the proof, we will again use Theorem A.1 to show that (84) is equivalent to $\mathbb{V}$ given in the theorem. To this end, define

$$F[f](x, s, t) := \frac{1}{2\varphi(x) t^{k-1}} \left(K_{k+1}^{(2)}(x, t(1-s)) + \beta(x) S_2^{(2)}(x, t(1-s)) \right). \quad (89)$$

Due to (16) and the induction hypothesis,

$$\lim_{t \to \infty} \frac{1}{2\varphi(x) t^{k-1}} K_{k+1}^{(2)}(x, t(1-s))$$

$$= \frac{(1-s)^{k-1}}{2\varphi(x)} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1! k_2!} \psi''(x, 0+) \lim_{t \to \infty} \frac{\mathrm{T}_{t(1-s)}^{(k_1)} f(x)}{(t(1-s))^{k_1 - 1}} \frac{\mathrm{T}_{t(1-s)}^{(k_2)} f(x)}{(t(1-s))^{k_2 - 1}}$$

$$= (1-s)^{k-1} \varphi(x) \sum_{\{k_1, k_2\}^+} (k+1)! 2^{-k} \psi''(x, 0+) \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1}$$

$$= k(1-s)^{k-1} \varphi(x)(k+1)! 2^{-k} \psi''(x, 0+) \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1},$$

where the last equality holds because the total number of ways of splitting one set of size $k+1$ into two non empty sets is equal to k. To obtain the limit for $S_{k+1}^{(2)}$, we use (16), the induction hypothesis, dominated convergence and linearity to obtain

$$\lim_{t \to \infty} \frac{S_{k+1}^{(2)}(x, t(1-s))}{2\varphi(x) t^{k-1}}$$

$$= \frac{(1-s)^{k-1}}{2\varphi(x)} \int_{M(E)^\circ} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1! k_2!} \lim_{t \to \infty} \frac{\langle \mathrm{T}_{t(1-s)}^{(k_1)}[f], \nu \rangle}{(t(1-s))^{k_1 - 1}} \frac{\langle \mathrm{T}_{t(1-s)}^{(k_2)}[f], \nu \rangle}{(t(1-s))^{k_2 - 1}} \Gamma(x, d\nu)$$

$$= \frac{(1-s)^{k-1}}{2^k \varphi(x)} \int_{M(E)^\circ} \sum_{\{k_1, k_2\}^+} (k+1)! \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \langle \varphi, \nu \rangle^2 \Gamma(x, d\nu)$$

$$= \frac{k(1-s)^{k-1}}{2^k \varphi(x)} (k+1)! \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \int_{M(E)^\circ} \langle \varphi, \nu \rangle^2 \Gamma(x, d\nu).$$

Combining these two limits, we get that

$$
\begin{aligned}
F[f](x, s) &:= \lim_{t \to \infty} F[f](x, s, t) \\
&= \frac{k(1-s)^{k-1}}{\varphi(x)} \frac{(k+1)!}{2^k} \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \mathbb{V}[\varphi](x).
\end{aligned}
$$

To complete the proof it remains to verify assumptions (A.1) and (A.2) in order to apply Theorem A.1 to (84). By now the reader will be familiar with the arguments required to check these assumptions and thus, we exclude the details. Hence, it follows that

$$
\begin{aligned}
\lim_{t \to \infty} \frac{1}{\varphi(x) t^k} \mathrm{T}^{(k+1)}[f](x) &= \frac{(k+1)!}{2^k} \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \\
&\quad \int_0^1 k(1-s)^{k-1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle \, ds \\
&= \frac{(k+1)!}{2^k} \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^k,
\end{aligned}
$$

where the limit is uniform in $x \in E$. Moreover, $\sup_{t \geq 0, x \in E} \mathrm{T}^{(k+1)}[f](x)/\varphi(x) t^k < \infty$.

3.4 Proofs for moments in the non-critical cases

In this section we present the main ideas behind the proof of Theorems 2 and 3. The methods follow a similar reasoning to the critical case and the details are left to the reader. The base case is given by the Perron Frobenius behaviour in (H1) for both sub and supercritical cases. Thus, we assume the result for $k-1$ and proceed to give the outline of the inductive step of the argument.

Proof of Theorem 2 (supercritical case) The main difference here, compared to the critical case, is that all the terms in $R_k(x, t)$ will survive after the normalisation $\mathrm{e}^{-\lambda k t}$ since the exponential term shares across the product. From the evolution equation (77) and the definition of L_k we have that

$$
\begin{aligned}
&|\varphi(x)^{-1} \mathrm{e}^{-\lambda k t} \mathrm{T}_t^{(k)}[f](x) - k! \langle f, \tilde{\varphi} \rangle^k L_k(x)| \\
&\leq \Big| \varphi(x)^{-1} \mathrm{e}^{-\lambda k t} (-1)^{k+1} R_k(x, t) \\
&\quad - \langle f, \tilde{\varphi} \rangle^k \sum_{\{m_1, \ldots, m_{k-1}\}} \frac{k!}{m_1! \ldots m_{k-1}!} (m_1 + \cdots + m_{k-1} - 1)! \varphi(x)^{m_1 + \cdots + m_{k-1} - 1} \\
&\quad \prod_{j=1}^{k-1} (-L_j(x))^{m_j} \Big| \\
&\quad + \Big| \varphi(x)^{-1} \mathrm{e}^{-\lambda k t} (-1)^k \int_0^t T_s[U_k(\cdot, t-s)](x) ds - \frac{k! \langle f, \tilde{\varphi} \rangle^k}{\lambda(k-1)} \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle \Big|.
\end{aligned}
$$

The first terms in the right hand side goes to zero uniformly since

$$\frac{e^{-\lambda kt}}{\varphi(x)}(-1)^{k+1}R_k(x,t) = \sum_{\{m_1,\ldots,m_{k-1}\}_k} \frac{k!}{m_1!\ldots m_{k-1}!}(-1)^{m_1+\cdots+m_{k-1}}$$

$$\times (m_1+\cdots+m_{k-1}-1)! \prod_{j=1}^{k-1} \left(\frac{e^{-\lambda jt}T_t^{(j)}[f](x)}{\varphi(x)j!}\right)^{m_j} \varphi(x)^{m_1+\cdots+m_{k-1}-1},$$

and the induction hypothesis implies that

$$\lim_{t\to\infty} \frac{e^{-\lambda kt}}{\varphi(x)}(-1)^{k+1}R_k(x,t) = \langle f,\tilde{\varphi}\rangle^k \sum_{\{m_1,\ldots,m_{k-1}\}_k} \frac{k!}{m_1!\ldots m_{k-1}!}$$

$$\times (m_1+\cdots+m_{k-1}-1)!\varphi(x)^{m_1+\cdots+m_{k-1}-1} \prod_{j=1}^{k-1}(-L_j(x))^{m_j}. \qquad (90)$$

For the second term, define for $k \geq 2$

$$I_k(x,t) := \int_0^t T_s[U_k(\cdot,t-s)](x)\mathrm{d}s = (-1)^k T_t^{(k)}[f](x) + R_k(x,t).$$

We will use induction to prove that for $k \geq 2$

$$\lim_{t\to\infty} \sup_{x\in E, f\in B_1^+(E)} \left| \frac{e^{-\lambda kt}}{\varphi(x)}(-1)^k I_k(x,t) - \frac{k!\langle f,\tilde{\varphi}\rangle^k}{\lambda(k-1)}\langle \mathbb{V}_k[\varphi],\tilde{\varphi}\rangle \right| = 0, \qquad (91)$$

which will complete the proof of the theorem.

First notice that for any $k \geq 2$ due to a change of variable, we have that

$$\frac{e^{-\lambda kt}}{\varphi(x)}(-1)^k I_k(x,t)$$

$$= t\int_0^1 e^{-\lambda(k-1)ut} \frac{e^{-\lambda ut}}{\varphi(x)} T_{ut}\left[(-1)^k e^{-\lambda kt(1-u)}U_k(\cdot,t(1-u))\right](x)du,$$

where

$$(-1)^k e^{-\lambda kt}U_k(x,t) = (-1)^k e^{-\lambda kt}K_k(x,t) + \beta(x)(-1)^k e^{-\lambda kt}S_k(x,t).$$

Recalling the definitions of K_k and S_k given in (74) and (75) respectively, and using the fact that $I_j(x,t) = -((-1)^{j+1}T_t[f](x) - R_j(x,t))$ for $j = 2,\ldots,k-1$, after sharing the exponential term across the products, we obtain

$$(-1)^k e^{-\lambda kt}K_k(x,t) = \sum_{\{m_1,\ldots,m_{k-1}\}_k} \frac{k!}{m_1!\ldots m_{k-1}!}$$

$$\psi^{(m_1+\cdots+m_{k-1})}(x,0+)\left(-e^{-\lambda t}\,T_t\,[f]\,(x)\right)^{m_1}$$

$$\prod_{j=2}^{k-1}\left(-\frac{1}{j!}(-1)^j e^{-\lambda j t} I_j(x,t)\right)^{m_j}$$

and

$$(-1)^k e^{-\lambda k t}\,S_k(x,t)$$

$$=\int_{M(E)^\circ}\sum_{\{m_1,\ldots,m_{k-1}\}_k}\frac{k!}{m_1!\ldots m_{k-1}!}(-1)^{m_1+\cdots+m_{k-1}}\left(-e^{-\lambda t}\,T_t\,[f]\,,v\right)^{m_1}$$

$$\prod_{j=2}^{k-1}\frac{1}{j!}\left(-(-1)^j e^{-\lambda j t} I_j(\cdot,t),v\right)^{m_j}\Gamma(x,dv).$$

From these expressions and the definition of $\mathbb{V}_2$, the case $k=2$ follows easily. Now we assume (91) holds for $\ell=1,\ldots,k-1$, then similarly to the branching Markov process case it follows that

$$\lim_{t\to\infty}\sup_{x\in E,\,f\in B_1^+(E)}\left|\frac{e^{-\lambda k t}}{\varphi(x)}(-1)^k I_k-\frac{k!\,\langle f,\tilde\varphi\rangle^k}{\lambda(k-1)}\,\langle\mathbb{V}_k\,[\varphi]\,,\tilde\varphi\rangle\right|$$

$$\leq\lim_{t\to\infty}\sup_{x\in E,\,f\in B_1^+(E)}$$

$$t\left|\int_0^1 e^{-\lambda(k-1)ut}\left(\frac{e^{-\lambda u t}}{\varphi(x)}\,T_{ut}\,[\varphi F(\cdot,u,t)]-k!\,\langle f,\tilde\varphi\rangle^k\,\langle\mathbb{V}_k\,[\varphi]\,,\tilde\varphi\rangle\right)du\right|,$$

$$(92)$$

where we have defined

$$F[f](x,u,t)$$

$$=\varphi(x)^{-1}\left((-1)^k e^{-\lambda k t(1-u)}K_k(x,t(1-u))+\beta(x)(-1)^k e^{-\lambda k t(1-u)}S_k(x,t(1-u))\right)$$

$$=\frac{1}{\varphi(x)}\sum_{\{m_1,\ldots,m_{k-1}\}_k}\frac{k!}{m_1!\ldots m_{k-1}!}$$

$$\times\left[\psi^{(m_1+\cdots+m_{k-1})}(-\varphi)^{m_1+\cdots+m_{k-1}}\left(\frac{e^{-\lambda t(1-u)}T_{t(1-u)}\,[f]\,(x)}{\varphi(x)}\right)^{m_1}\right.$$

$$\prod_{j=2}^{k-1}\left(\frac{(-1)^j e^{-\lambda j t(1-u)}I_j(x,t(1-u))}{\varphi(x)j!}\right)^{m_j}$$

$$+\beta(x)\int_{M(E)^\circ}\left\langle\frac{e^{\lambda t(1-u)}T_{t(1-u)}\,[f]}{\varphi}\varphi,v\right\rangle^{m_1}\prod_{j=2}^{k-1}\frac{1}{j!}$$

$$\left\langle \frac{(-1)^j e^{-\lambda j t(1-u)} I_j(\cdot, t(1-u))}{\varphi} \varphi, v \right\rangle^{m_j} \Gamma(x, dv) \Bigg].$$

It is easy to see that, pointwise in $x \in E$ and for $u \in (0, 1)$, using the induction hypothesis for I_k and the assumed Perron Frobenius behaviour (H1) for $T_{t(1-u)}[f](x)$ we have

$$F[f](x, u) := \lim_{t \to \infty} F[f](x, u, t)$$

$$= \frac{1}{\varphi(x)} \sum_{\{m_1, \ldots, m_{k-1}\}_k} \frac{k!}{m_1! \ldots m_{k-1}!} \langle f, \tilde\varphi \rangle^{m_1} \prod_{j=2}^{k-1} \left(\frac{\langle f, \tilde\varphi \rangle^j \langle \mathbb{V}_j[\varphi], \tilde\varphi \rangle}{\lambda(j-1)} \right)^{m_j}$$

$$\times \Bigg[\psi^{(m_1 + \cdots + m_{k-1})}(x, 0+)(-\varphi(x))^{m_1 + \cdots + m_{k-1}}$$

$$+ \beta(x) \int_{M(E)^\circ} \langle \varphi, v \rangle^{m_1 + \cdots + m_{k-1}} \Gamma(x, dv) \Bigg]$$

$$= \frac{k!}{\varphi(x)} \langle f, \tilde\varphi \rangle^k \mathbb{V}_k[\varphi](x).$$

We again verify that the conditions (A.1) and (A.2) hold using the induction hypothesis and (H2). To complete the proof of (91), we again proceed along the same lines as in the branching Markov processes setting. Similar arguments to those given in the proof of Theorem A.1

$$\sup_{\substack{x \in E, u \in [0,1], \\ f \in B_1^+(E), t \geq 0}} \left| \frac{e^{-\lambda u t}}{\varphi(x)} T_{ut}[\varphi F[f](x, u, t)] - k! \langle f, \tilde\varphi \rangle^k \mathbb{V}_k[\varphi] \right| < \infty,$$

it follows that the remainder of the integral in (92) can be bounded by $\varepsilon(1 - e^{-\lambda(k-1)t})/\lambda(k-1)$, which can be bounded by ε. Combining this with (90) we get the desired result. $\qquad \square$

Proof of Theorem 3 (subcritical case) We now outline the proof for the subcritical case. Again we use an inductive argument. The case $k = 1$ follows from (H1) and the fact that $\langle \varphi, \tilde\varphi \rangle = 1$. Now assume the result to be true for $\ell = 1, \ldots, k - 1$. We first note first that the term $R_k(x, t)$ in (77) vanishes in the limit after the normalisation $e^{-\lambda t}$. To see this, note from (73) that

$$\left| \frac{e^{-\lambda t}}{\varphi(x)} R_k(x, t) \right| \leq \sum_{\{m_1, \ldots, m_{k-1}\}_k} c(m_1, \ldots, m_{k-1}) \prod_{j=1}^{k-1} \left| \frac{e^{-\lambda t} T_t^{(j)}[f](x)}{\varphi(x) j!} \right|^{m_j}$$

$$\times \varphi(x)^{m_1 + \cdots + m_{k-1}} e^{\lambda(m_1 + \cdots + m_{k-1} - 1)t},$$

where $c(m_1, \ldots, m_{k-1})$ is a constant depending only on $m_1, \ldots, m_{k-1}$. Since each of the terms in the product is bounded, $\lambda < 0$, and $m_1 + \cdots + m_{k-1} > 1$ for any

partition, the limit of the right-hand side above is zero. Using this, along with the induction hypothesis, assumption (H1) and the evolution equation (77) we see that

$$
\lim_{t\to\infty} \sup_{x\in E,\, f\in B_1^+(E)} \left| \varphi(x)^{-1} e^{-\lambda t} T_t^{(k)}[f](x) - k! \langle f, \tilde{\varphi} \rangle^k \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle \right|
$$

$$
= \lim_{t\to\infty} \sup_{x\in E,\, f\in B_1^+(E)} \left| \frac{e^{-\lambda t}}{\varphi(x)} \int_0^t \mathrm{T}_s \left[U_k^*(\cdot, t-s) \right](x)\mathrm{d}s - k! \langle f, \tilde{\varphi} \rangle^k \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle \right|,
$$

(93)

where

$$
U_k^*(x,t) = \sum_{\{m_1 \ldots, m_{k-1}\}_k} \frac{k!}{m_1! \ldots m_{k-1}!} \left[\psi^{(m_1+\cdots+m_{k-1})}(x,0+) \prod_{j=1}^{k-1} \left(\frac{-T_t^{(j)}[f](x)}{j!} \right)^{m_j} \right.
$$

$$
\left. + \beta(x) \int_{M(E)^\circ} \prod_{j=1}^{k-1} \left(\frac{1}{j!} \left\langle T_t^{(j)}[f], \nu \right\rangle \right)^{m_j} \Gamma(x,\mathrm{d}\nu) \right].
$$

Then, similar calculations to those above yield

$$
\frac{e^{-\lambda t}}{\varphi(x)} \int_0^t \mathrm{T}_s \left[U_k^*(\cdot, t-s) \right](x)\mathrm{d}s
$$

$$
= t \sum_{\{m_1, \ldots, m_{k-1}\}_k} \frac{k!}{m_1! \ldots m_{k-1}!} \int_0^1 e^{-\lambda t(1-u)(1-m_1-\cdots-m_{k-1}))}
$$

$$
\frac{e^{-\lambda t u}}{\varphi(x)} \mathrm{T}_{tu} \left[\psi^{(m_1+\cdots+m_{k-1})}(\cdot, 0+)(-\varphi(\cdot))^{m_1+\cdots+m_{k-1}} \prod_{j=1}^{k-1} \left(\frac{e^{-\lambda t(1-u)}T_{t(1-u)}^{(j)}[f](\cdot)}{\varphi(\cdot)j!} \right)^{m_j} \right.
$$

$$
\left. + \beta(\cdot) \int_{M(E)^\circ} \prod_{j=1}^{k-1} \left\langle \varphi \frac{e^{-\lambda t(1-u)}}{\varphi j!} T_t^{(j)}[f], \nu \right\rangle^{m_j} \Gamma(x,\mathrm{d}\nu) \right](x)\mathrm{d}u.
$$

To finish the proof, we use induction hypothesis to deal with the lower order moments, dominated convergence and (H1) to deal with the limit of $\mathrm{T}_{t(1-u)}$. Similarly to the last part of the proof of Theorem 2 we get that (93) is bounded as $t \to \infty$ by $\varepsilon(1 - e^{-\lambda t(1-m_1-\cdots-m_k)})/\lambda(1 - m_1 - \cdots - m_k)$, which is bounded by ε since $m_1 + \cdots + m_{k-1} > 1$. We once more leave the details of the rest of the proof to the reader. $\square$

3.5 Proofs for occupation moments, Theorems 4, 5 and 6

Given the proofs we have now seen for the branching particle setting for these three theorems, as well as the proofs of Theorem 1, 2 and 3, we mention only that a similar

calculation to the one presented in Theorem 7 tells us that

$$
\mathrm{M}_t^{(k)}[g](x) = (-1)^{k+1}\tilde{R}_k(x,t) + (-1)^k \int_0^t \mathrm{T}_s\left[\tilde{U}_k(\cdot, t-s)\right](x)\mathrm{d}s
$$
$$
- k \int_0^t \mathrm{T}_s\left[g[\mathrm{M}_{t-s}^{(k-1)}[g] + (-1)^{k-1}\tilde{R}_{k-1}(\cdot, t-s)]\right](x)\mathrm{d}s,
$$

where

$$
\tilde{U}_k(x,t) = \tilde{K}_k(x,t) + \beta(x)\tilde{S}_k(x,t)
$$

and $\tilde{R}$, $\tilde{K}$ and $\tilde{S}$ are defined as R, K and S in (73), (74), (75), respectively, albeit replacing $\mathrm{T}^{(j)}$ by $\mathrm{M}^{(j)}$. From here we can consider the claimed asymptotics using the the inductive reasoning as in the proof of Theorems 1, 2 and 3, respectively. $\qquad\square$

Appendix

We first state two fundamental combinatorial results for complex derivatives, the classical Faà di Bruno and Leibniz rules. In both cases, for a sufficiently smooth function g on $\mathbb{R}$, we will denote by $g^{(k)}$ by its k-th derivative.

Lemma A.1 (Faà di Bruno rule) *Let f and g k-times continuously differentiable functions on $\mathbb{R}$. Then the k-th derivative is given by the following formula*

$$
\frac{\mathrm{d}^k}{\mathrm{d}x^k}f(g(x)) = \sum_{\{m_1,\ldots,m_k\}_k} \frac{k!}{m_1!\ldots m_k!} f^{(m_1+\cdots+m_k)}(g(x)) \prod_{j=1}^k \left(\frac{g^{(j)}(x)}{j!}\right)^{m_j},
$$

where the sum goes over the set $\{m_1, \ldots, m_k\}_k$ of non-negative integers such that

$$
m_1 + 2m_2 + \cdots + km_k = k.
$$

Lemma A.2 (Leibniz rule) *Suppose $g_1, \ldots, g_m$ are k-times continuously differentiable functions on $\mathbb{R}$, for $k \geq 1$. Then*

$$
\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left(\prod_{i=1}^m g_i(x)\right) = \sum_{k_1+\cdots+k_m=k} \binom{k}{k_1, \ldots, k_m} \prod_{\ell=1}^m g_\ell^{(k_\ell)}(x),
$$

where the sum is taken over all non-negative integers $k_1, \ldots, k_m$ such that $\sum_{i=1}^{m} k_i = k$.

The third result of the appendix is a general ergodic limit theorem which is key to the moment convergence. We will only state the result in the critical case, since we will only apply it in the proof of Theorem 1, however, the result can easily be extended to the non-critical case by including the normalisation $e^{-\lambda ut}$ in the first integral.

In order to state it, let us introduce a class of functions $\mathcal{C}$ on $B_1^+(E) \times E \times [0, 1] \times [0, \infty)$ such that F belongs to class $\mathcal{C}$ if

$$F[g](x, s) := \lim_{t \to \infty} F[g](x, s, t), \qquad g \in B_1^+(E), x \in E, s \in [0, 1],$$

exists,

$$\sup_{x \in E, s \in [0,1], g \in B_1^+(E)} |\varphi F[g](x, s)| < \infty, \tag{A.1}$$

and

$$\lim_{t \to \infty} \sup_{x \in E, s \in [0,1], g \in B_1^+(E)} \varphi(x) |F[g](x, s) - F[g](x, s, t)| = 0. \tag{A.2}$$

Theorem A.1 *Assume (H1) holds, $\lambda = 0$ and that $F \in \mathcal{C}$. Define*

$$\Xi_t = \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \int_0^1 T_{ut}[\varphi F[g](\cdot, u, t)](x)\mathrm{d}u - \int_0^1 \langle F[g](\cdot, u), \varphi\tilde{\varphi}\rangle \mathrm{d}u \right|,$$

$$t \geq 0.$$

Then

$$\sup_{t \geq 0} \Xi_t < \infty \text{ and } \lim_{t \to \infty} \Xi_t = 0. \tag{A.3}$$

Proof We will show that

$$\lim_{t \to \infty} \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} T_{ut}[\varphi F[g](\cdot, u, t)](x) - \langle \tilde{\varphi}\varphi, F[g](\cdot, u)\rangle \right| = 0,$$

since then

$$\lim_{t \to \infty} \sup_{x \in E, g \in B_1^+(E)} \left| \int_0^1 \frac{1}{\varphi(x)} T_{ut}[\varphi F[g](\cdot, u, t)](x)\mathrm{d}u - \int_0^1 \langle \tilde{\varphi}\varphi, F[g](\cdot, u)\rangle \mathrm{d}u \right|$$

$$\leq \int_0^1 \lim_{t \to \infty} \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} T_{ut}[\varphi F[g](\cdot, u, t)](x) - \langle \tilde{\varphi}\varphi, F[g](\cdot, u)\rangle \right| \mathrm{d}u = 0.$$

First note that,

$$\left| \frac{1}{\varphi(x)} \mathtt{T}_{ut}[\varphi F[g](\cdot, u, t)](x) - \langle \tilde{\varphi}\varphi, F[g](\cdot, u) \rangle \right|$$

$$\leq \frac{1}{\varphi(x)} \mathtt{T}_{ut}[|\varphi F[g](\cdot, u, t) - \varphi F[g](\cdot, u)|](x)$$

$$+ \left| \frac{1}{\varphi(x)} \mathtt{T}_{ut}[\varphi F[g](\cdot, u)](x) - \langle \tilde{\varphi}, \varphi F[g](\cdot, u) \rangle \right|$$

Due to assumption (A.2), for t sufficiently large, the first term on the right-hand side above can be controlled by $\varphi^{-1}(x) \mathtt{T}_{ut}[\varepsilon](x)$. Combining this with the above inequality yields

$$\sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \mathtt{T}_{ut}[\varphi F[g](\cdot, u, t)](x) - \langle \tilde{\varphi}\varphi, F[g](\cdot, u) \rangle \right|$$

$$\leq \sup_{x \in E} \left| \varphi^{-1}(x) \mathtt{T}_{ut}[\varepsilon](x) - \langle \tilde{\varphi}, \varepsilon \rangle \right| + \varepsilon \|\tilde{\varphi}\|_1$$

$$+ \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \mathtt{T}_{ut}[\varphi F[g](\cdot, u)](x) - \langle \tilde{\varphi}, \varphi F[g](\cdot, u) \rangle \right|. \qquad (A.4)$$

We note that (A.1) and the first (resp. second) statement of (14) in (H1), together with dominated convergence, immediately imply that the first (resp. second) statement in (A.3) holds. □

References

1. Chen, Z.-Q., Ren, Y.-X., Yang, T.: Law of large numbers for branching symmetric Hunt processes with measure-valued branching rates. J. Theoret. Probab. **30**(3), 898–931 (2017)
2. Chen, Z.-Q., Ren, Y.-X., Yang, T.: Skeleton decomposition and law of large numbers for supercritical superprocesses. Acta Appl. Math. **159**, 225–285 (2019)
3. Cox, A.M.G., Harris, S.C., Horton, E.L., Kyprianou, A.E.: Multi-species neutron transport equation. J. Stat. Phys. **176**(2), 425–455 (2019)
4. Crump, K.S., Mode, C.J.: A general age-dependent branching process. i. J. Math. Anal. Appl. **24**(3), 494–508 (1968)
5. Crump, K.S., Mode, C.J.: A general age-dependent branching process. ii. J. Math. Anal. Appl. **25**(1), 8–17 (1969)
6. Dawson, D.A.: Measure-valued Markov processes. In: *École d'Été de Probabilités de Saint-Flour XXI—1991*, volume 1541 of Lecture Notes in Mathematics, pp. 1–260. Springer, Berlin (1993)
7. Dawson, D.A., Gorostiza, L.G., Li, Z.: Nonlocal branching superprocesses and some related models. Acta Appl. Math. **74**(1), 93–112 (2002)
8. Dumonteil, E., Bahran, R., Cutler, T., Dechenaux, B., Grove, T., Hutchinson, J., McKenzie, G., McSpaden, A., Monange, W., Nelson, M., Thompson, N., Zoia, A.: Patchy nuclear chain reactions. Commun. Phys. **4**(1), 1–10 (2021)
9. Dumonteil, E., Malvagi, F., Zoia, A., Mazzolo, A., Artusio, D., Dieudonné, C., De Mulatier, C.: Particle clustering in Monte Carlo criticality simulations. Ann. Nucl. Energy **63**, 612–618 (2014)
10. Dumonteil, E., Mazzolo, A.: Residence times of branching diffusion processes. Phys. Rev. E **94**, 012131 (2016)

11. Durham, S.D.: Limit theorems for a general critical branching process. J. Appl. Probab. **8**(1), 1–16 (1971)
12. Dynkin, E.B.: Branching particle systems and superprocesses. Ann. Probab. **19**(3), 1157–1194 (1991)
13. Dynkin, E.B.: Diffusions, Superdiffusions, and Partial Differential Equations, vol. 50. American Mathematical Society, Providence, RI (2002)
14. Eckhoff, M., Kyprianou, A.E., Winkel, M.: Spines, skeletons and the strong law of large numbers for superdiffusions. Ann. Probab. **43**(5), 2545–2610 (2015)
15. Engländer, J.: Spatial Branching in Random Environments and with Interaction. Advanced Series on Statistical Science & Applied Probability, vol. 20. World Scientific Publishing, Hackensack, NJ (2015)
16. Engländer, J., Harris, S.C., Kyprianou, A.E.: Strong law of large numbers for branching diffusions. Ann. Inst. Henri Poincaré Probab. Stat. **46**(1), 279–298 (2010)
17. Etheridge, A.M.: An Introduction to Superprocesses. University Lecture Series, vol. 20. American Mathematical Society, Providence, RI (2000)
18. Fleischman, J.: Limiting distributions for branching random fields. Trans. Am. Math. Soc. **239**, 353–389 (1978)
19. Harris, S.C., Horton, E., Kyprianou, A.E.: Stochastic methods for the neutron transport equation II: almost sure growth. Ann. Appl. Probab. **30**(6), 2815–2845 (2020)
20. Harris, S.C., Horton, E., Kyprianou, A.E., Wang, M.: Yaglom limit for critical neutron transport (2021)
21. Harris, S.C., Roberts, M.I.: The many-to-few lemma and multiple spines. Ann. Inst. Henri Poincaré Probab. Stat. **53**(1), 226–242 (2017)
22. Horton, E., Kyprianou, A.E., Villemonais, D.: Stochastic methods for the neutron transport equation I: linear semigroup asymptotics. Ann. Appl. Probab. **30**(6), 2573–2612 (2020)
23. Iscoe, I.: On the supports of measure-valued critical branching Brownian motion. Ann. Probab. **16**(1), 200–221 (1988)
24. Klenke, A.: Multiple scale analysis of clusters in spatial branching models. Ann. Probab. **25**(4), 1670–1711 (1997)
25. Koralov, L., Molchanov, S.: Structure of population inside propagating front. J. Math. Sci. **189**(4), 637–658 (2013)
26. Li, Z.: Measure Valued Branching Markov Processes. Springer-Verlag, Berlin (2011)
27. Palau, S., Yang, T.: Law of large numbers for supercritical superprocesses with non-local branching. Stoch. Process. Appl. **130**(2), 1074–1102 (2020)
28. Powell, E.: An invariance principle for branching diffusions in bounded domains. Probab. Theory Relat. Fields **173**(3–4), 999–1062 (2019)
29. Ren, Y.-X., Song, R., Sun, Z.: Spine decompositions and limit theorems for a class of critical superprocesses. Acta Appl. Math. **165**, 91–131 (2020)
30. Sevast'janov, B.A.: The extinction conditions for branching processes with diffusion. Teor. Verojatnost. i Primenen. **6**, 276–286 (1961)
31. Sevast'yanov, B.A.: Branching stochastic processes for particles diffusing in a bounded domain with absorbing boundaries. Teor. Veroyatnost. i Primenen. **3**, 121–136 (1958)
32. Skorohod, A.V.: Branching diffusion processes. Teor. Verojatnost. i Primenen. **9**, 492–497 (1964)
33. Sutton, T.M., Mittal, A.: Neutron clustering in Monte Carlo iterated-source calculations. Nucl. Eng. Technol. **49**(6), 1211–1218 (2017)
34. Watanabe, S.: A limit theorem of branching processes and continuous state branching processes. J. Math. Kyoto Univ. **8**, 141–167 (1968)
35. Zel'Dovich, Y.A.B., Molchanov, S.A., Ruzmaĭkin, A.A., Sokolov, D.D.: Intermittency in random media. Soviet Phys. Uspekhi **30**(5), 353 (1987)